

Analysis of the functions and  
interactions of the RAD51  
paralogues in *Trypanosoma brucei*.

Rachel Pamela Dobson

University of Glasgow  
Wellcome Centre for Molecular Parasitology  
Glasgow Biomedical Research Centre

Submitted in fulfilment of the requirements for the  
Degree of Doctor of Philosophy

June, 2009

## **Abstract**

*Trypanosoma brucei* evades host acquired immunity by antigenic variation, involving periodic switches in the variant surface glycoprotein (VSG) coat. DNA recombination is critical in this process and a key component of homologous recombination, RAD51, also plays a role in VSG switching. *T. brucei* encodes four proteins distantly related to RAD51; termed RAD51 paralogues, named RAD51-3, RAD51-4, RAD51-5 and RAD51-6. Two of these RAD51 paralogues, RAD51-3 and RAD51-5, have been shown to function in homologous recombination, DNA repair and RAD51 re-localization into foci following DNA damage. Surprisingly, however, only RAD51-3 appears to act in VSG switching. To examine the functions of all the RAD51 paralogues in *T. brucei*, reverse genetics has been used to generate mutants of the two remaining unstudied paralogues, RAD51-4 and RAD51-6. Phenotypic analysis of both mutant cell lines indicates that these factors also play critical roles in RAD51-directed recombination and repair, and both influence VSG switching in the parasite. As homozygous mutant cell lines of RAD51 and the RAD51 paralogues were available, it was possible to comprehensively compare the phenotypes of *rad51* *-/-* with *rad51-3* *-/-*, *rad51-4* *-/-*, *rad51-5* *-/-*, and *rad51-6* *-/-*. From these results it was observed that the phenotypes of the *rad51* paralogue mutants are broadly equivalent, with two exceptions. Firstly, as mentioned above, RAD51-5 does not function in VSG switching, and RAD51-4 and RAD51-6 may not have direct roles in VSG switching. Secondly, *rad51-4* *-/-* mutants are less sensitive to the DNA damaging agent phleomycin and a higher percentage of *rad51-4* *-/-* cells form DNA damage-induced RAD51 foci compared with the other homozygous mutant cell lines. These results may imply that RAD51-4 and RAD51-5 have a less central role in RAD51-directed DNA repair or that their functions can be performed by other factors. In addition, the physical interactions of all the RAD51 paralogues were examined. It was found by yeast two-hybrid and co-immunoprecipitation that they form at least two complexes, and probably function in sub-complexes in a similar manner to the Rad51 paralogues of higher eukaryotes. These analyses shed light on the evolution and role of eukaryotic RAD51 paralogues in DNA recombination and repair in general, as well as the contribution that recombination makes to antigenic variation in *T. brucei*.

## **Table of Contents**

Abstract .....	ii
List of Contents.....	iii
List of Tables.....	viii
List of Figures.....	ix
Preface.....	xv
Acknowledgement.....	xvi
Authors Declaration.....	xvii
Abbreviations.....	xvi

ii

Chapter 1. Introduction.....	1
1.1 General introduction to <i>Trypanosoma brucei</i> .....	2
1.1.1 Phylogeny of the African trypanosome .....	2
1.1.2 Epidemiology and Life cycle .....	3
1.1.3 Trypanosomiasis: symptoms & treatment. ....	4
1.1.4 Use of <i>T. brucei</i> as a model organism?.....	5
1.1.5 Genome structure and gene expression by poly-cistronic transcription .....	6
1.2 Antigenic Variation.....	8
1.2.1 The role of Variant Surface Glycoproteins (VSG) in antigenic variation .....	9
1.2.2 Mechanisms of VSG switching .....	11
1.3 DNA repair.....	17
1.4 End-Joining Double Strand Break repair.....	20
1.4.1 Non-Homologous End Joining (NHEJ).....	20
1.4.2 Micro-homology Mediated End-Joining (MMEJ).....	22
1.5 Homologous Recombination (HR) .....	23
1.5.1 Mechanism of Homologous Recombination: Gene conversion .....	23
1.5.2 Break induced replication .....	28
1.5.3 Single strand annealing .....	28
1.5.4 Other roles of homologous recombination .....	28
1.6 The role of the HR recombinases.....	32
1.6.1 Phylogeny of the recombinases and their role in strand exchange .....	32
1.6.2 Phenotypes of recombinase mutants.....	35
1.7 The roles of the eukaryotic RAD51 paralogues.....	35
1.7.1 Identified Rad51 paralogues .....	36
1.7.2 Phenotypes of <i>rad51</i> paralogue mutants in yeast .....	36
1.7.3 Interactions and roles of yeast RAD51 paralogues.....	37
1.7.4 Phenotypes of <i>rad51</i> paralogue mutants in vertebrate cells .....	38
1.7.5 Interactions and roles of Rad51 paralogues in vertebrate cells.....	39
1.7.6 Phenotypes of the <i>rad51-3</i> and <i>rad51-5</i> mutants in <i>T. brucei</i> cells.....	41
1.7.7 Some examples of the phenotypes of the Rad51 paralogue mutants in other eukaryotic cells .....	41
1.8 Aims of the project.....	42
1.8.1 Examining the role of <i>RAD51-4</i> and <i>RAD51-6</i> .....	42
1.8.2 Analysing the association between the RAD51-paralogues .....	43
1.8.3 A luciferase assay for <i>T. brucei</i> homologous recombination .....	43
Chapter 2. Materials and Methods.....	44
2.1 Cell culturing .....	45
2.1.1 Trypanosome growth .....	45

2.1.2	Transfecting of <i>T. brucei</i> .....	45
2.2	Isolation of DNA.....	46
2.2.1	Methods of DNA extraction.....	46
2.2.2	Phenol/Chloroform extraction and ethanol precipitation of DNA.....	46
2.3	RNA extraction .....	47
2.3.1	RNA extraction from <i>T. brucei</i> cells.....	47
2.3.2	Reverse Transcriptase-PCR (RT-PCR).....	47
2.4	Cloning procedure.....	48
2.4.1	Polymerase Chain Reaction (PCR).....	48
2.4.2	Restriction enzyme digestion of DNA .....	52
2.4.3	Gel electrophoresis and DNA fragment purification .....	52
2.4.4	DNA Ligation .....	52
2.4.5	Transformation into <i>E. coli</i> .....	53
2.4.6	Transformation into <i>S. cerevisiae</i> .....	53
2.4.7	TOPO TA cloning.....	54
2.5	Southern blotting.....	54
2.6	Protein analysis .....	55
2.6.1	Protein extraction from <i>S. cerevisiae</i> cells.....	55
2.6.2	Protein extraction from <i>T. brucei</i> cells .....	55
2.6.3	Electrophoresis of SDS-polyacrylamide gels and Western blotting.....	56
2.7	Purification of anti-RAD51 and the anti-RAD51 paralogue anti-sera.....	56
2.7.1	Anti-RAD51-5 purification from chicken egg yolks .....	57
2.7.2	Affinity Purification of anti-RAD51 and the anti-RAD51 paralogue anti-sera .....	57
2.7.3	Acetone protein extraction for purification of anti-RAD51-4 anti-serum .....	58
2.8	Analysing the knockout phenotypes .....	58
2.8.1	Growth curve and doubling times.....	58
2.8.2	DNA damage sensitivity .....	59
2.8.3	Transformation efficiency assay .....	59
2.8.4	VSG switching frequency and mechanism analysis .....	60
2.9	Immuno-localisation .....	62
2.9.1	Fixing <i>T. brucei</i> cells to slides .....	62
2.9.2	<i>in situ</i> fluorescence microscopy and DAPI staining of <i>T. brucei</i> cells.....	62
2.10	Assaying yeast two-hybrid interaction.....	63
2.10.1	To test for histidine auxotrophy .....	63
2.10.2	$\beta$ -galactosidase filter lift assay .....	63
2.10.3	Yeast $\beta$ -galactosidase assay kit.....	64
2.11	Co-immunoprecipitation .....	64
2.12	PTP purification of protein complexes .....	65
2.13	Bioluminescence detection .....	66
Chapter 3.	Role of two <i>RAD51</i> paralogues in <i>T. brucei</i> , <i>RAD51-4</i> and <i>RAD51-6</i> . .....	67
3.1	Introduction.....	68
3.2	Generation of the <i>RAD51-4</i> gene disruption mutant.....	69
3.2.1	Design of the <i>RAD51-4</i> knockout constructs.....	69
3.2.2	Generating the <i>RAD51-4</i> mutant cell lines .....	70
3.2.3	Confirmation of <i>RAD51-4</i> gene disruption mutants .....	71
3.3	Generation of <i>RAD51-4</i> re-expressor cell lines .....	75
3.3.1	Confirmation of the <i>RAD51-4</i> re-expressor cell lines by PCR and RT-PCR .....	76



3.3.2	Confirmation of the <i>RAD51-4</i> re-expressors by Western blot analysis .....	77
3.4	Phenotypic analysis of <i>RAD51-4</i> mutants and re-expressor cell lines.....	78
3.4.1	Analysis of <i>in vitro</i> growth rate of <i>RAD51-4</i> mutant cell lines .....	79
3.4.2	Analysis of DNA-damage sensitivity of <i>RAD51-4</i> mutant cell lines .....	81
3.4.3	Analysis of the recombination efficiency of <i>RAD51-4</i> mutant cell lines ....	88
3.4.4	Analysis of RAD51 foci formation in the <i>RAD51-4</i> mutant cell lines .....	90
3.4.5	Analysis of VSG switching of mutant cell lines.....	95
3.4.6	Analysis of VSG switching of <i>RAD51-4</i> mutant cell lines.....	97
3.5	Generation of the <i>RAD51-6</i> gene disruption mutants .....	101
3.5.1	Design of the <i>RAD51-6</i> knockout constructs.....	101
3.5.2	Generating the <i>RAD51-6</i> mutant cell lines .....	102
3.5.3	Confirmation of <i>RAD51-6</i> mutant cell lines .....	102
3.6	Generation of <i>RAD51-6</i> re-expressor cell lines .....	105
3.6.1	Confirmation of the <i>RAD51-6</i> re-expressor cell lines by RT-PCR.....	106
3.6.2	Confirmation of the <i>RAD51-6</i> re-expressors by Southern analysis .....	107
3.7	Phenotypic analysis of <i>RAD51-6</i> mutants and re-expressor cell lines.....	109
3.7.1	Analysis of <i>in vitro</i> growth rate of <i>RAD51-6</i> mutant cell lines .....	109
3.7.2	Analysis of DNA damage sensitivity of <i>RAD51-6</i> mutant cell lines.....	110
3.7.3	Analysis of recombination efficiency of <i>RAD51-6</i> mutant cell lines .....	117
3.7.4	Analysis of RAD51 foci formation in <i>RAD51-6</i> mutant cell lines .....	118
3.7.5	Analysis of VSG switching in <i>RAD51-6</i> mutant cell lines .....	121
3.8	Summary .....	123
Chapter 4.	Comparing the impact of gene knockout of <i>RAD51</i> and the <i>RAD51</i> paralogues in <i>T. brucei</i> .....	125
4.1	Introduction.....	126
4.2	Analysis of the <i>in vitro</i> growth rates of <i>RAD51</i> and <i>RAD51</i> paralogue mutants.....	129
4.3	Analysis of the DNA-damage sensitivity of <i>RAD51</i> and <i>RAD51</i> paralogue mutants.....	132
4.4	Analysis of the recombination efficiency of <i>RAD51</i> and the <i>RAD51</i> paralogue mutants .....	136
4.5	Analysis of the role of the RAD51 paralogues in RAD51 foci formation.....	138
4.6	Comparison of VSG switching in <i>RAD51</i> and the <i>RAD51</i> paralogue mutants .	143
4.7	Summary .....	145
Chapter 5.	Examining the associations between <i>T. brucei</i> RAD51 and the RAD51 paralogues .....	147
5.1	Introduction.....	148
5.2	Principles of the yeast two-hybrid system .....	149
5.3	Components of the yeast two-hybrid system .....	150
5.3.1	The role of LexA.....	150
5.3.2	The role of B42 .....	152
5.3.3	<i>S. cerevisiae</i> L40 strain .....	153
5.3.4	Independent testing of protein interactions.....	154
5.4	Results for yeast two-hybrid assays.....	156
5.4.1	Yeast two-hybrid analysis with LexA-RAD51 as the DNA-binding domain fusion.....	156
5.4.2	Yeast two-hybrid analysis with LexA-RAD51-3 as the DNA-binding domain fusion.....	160
5.4.3	Yeast two-hybrid analysis with LexA-RAD51-4 as the DNA-binding domain fusion.....	164

5.4.4	Yeast two-hybrid analysis with LexA-RAD51-5 as the DNA-binding domain fusion.....	171
5.4.5	Yeast two-hybrid analysis with LexA-RAD51-6 as the DNA-binding domain fusion.....	173
5.5	Introduction to the raising of anti-RAD51 and anti-RAD51 paralogue anti-sera .....	176
5.5.1	Anti-RAD51 anti-serum purification.....	178
5.5.2	Anti-RAD51-3 anti-serum purification.....	181
5.5.3	Anti-RAD51-4 anti-serum purification.....	183
5.5.4	Anti-RAD51-5 anti-serum purification.....	187
5.5.5	Anti-RAD51-6 anti-serum purification.....	190
5.6	Co-immunoprecipitation of the RAD51 paralogues .....	192
5.6.1	Principles of co-immunoprecipitation.....	192
5.6.2	Co-immunoprecipitation using the affinity purified anti-RAD51-3 anti-serum .....	193
5.6.3	Co-immunoprecipitation using the affinity purified anti-RAD51-4 anti-serum .....	195
5.7	PTP-tag purification of protein complexes .....	197
5.7.1	Principles of PTP-tag purifications.....	197
5.7.2	Generation of the PTP-tagged RAD51 paralogue constructs .....	199
5.7.3	Attempt to recover protein complexes by PTP purification .....	205
5.8	Possible co-expression of RAD51-3 and RAD51-4 .....	208
5.9	Summary .....	209
Chapter 6.	A luciferase assay to analyse homologous recombination in <i>T. brucei</i> .....	211
6.1	Introduction: Transformation efficiency to analyse homologous recombination rate .....	212
6.2	Principles of marker constructs to measure homologous recombination .....	214
6.3	Generation of a positive control expressing luciferase .....	216
6.4	A direct repeat luciferase construct to assay homologous recombination.....	219
6.4.1	Generation of the luciferase construct .....	219
6.4.2	Transformation of the construct into <i>T. brucei</i> cells.....	221
6.4.3	Assaying for homologous recombination using the luciferase construct with homologous regions as direct repeats .....	222
6.5	An inverted repeat luciferase construct to assay homologous recombination...	226
6.5.1	Generation of a luciferase construct with inverted homologous ends .....	226
6.6	Summary .....	227
Chapter 7.	Discussion .....	229
7.1	Introduction.....	230
7.2	Phylogeny of the <i>T. brucei</i> RAD51 paralogues .....	230
7.2.1	Identification of the <i>T. brucei</i> RAD51 paralogues.....	230
7.2.2	Varied number of Rad51 paralogues in different species.....	234
7.3	Roles of the <i>T. brucei</i> RAD51 paralogues .....	243
7.3.1	DNA repair and homologous recombination in <i>T. brucei</i> RAD51 paralogue mutants .....	243
7.3.2	RAD51 foci formation in <i>T. brucei</i> RAD51 paralogue mutants .....	244
7.3.3	VSG switching observed in the RAD51 paralogue mutants .....	246
7.4	Interactions between the RAD51 paralogues.....	248
7.5	Future goals.....	254
	Appendices.....	256

Appendix 1: <i>RAD51-4</i>	
ORF.....	256
Appendix 2: <i>RAD51-6</i>	
ORF.....	259
Appendix 3: Student's t-tests for the phenotypic analysis of the <i>rad51-4</i> <i>-/-</i> and <i>rad51-6</i> <i>-/-</i> mutants.....	262
Appendix 4: Accession numbers of the <i>RAD51</i> -related genes.....	269
Bibliography.....	272

## **List of Tables**

Table 2-1: Description of the primers used in the experiments described in this thesis.....	49
Table 2-2: Description of the vectors used in the experiments described in this thesis. ....	53
Table 3-1: <i>in vitro</i> doubling times for <i>RAD51-4</i> mutant cell lines. ....	80
Table 3-2 : Statistical analysis of the <i>in vitro</i> doubling times for <i>RAD51-4</i> mutants. ....	80
Table 3-3: Statistical analysis of the clonal survival of <i>RAD51-4</i> mutants when treated with 0.0003 % methyl methanesulfonate (MMS).....	83
Table 3-4: Statistical analysis of the clonal survival of <i>RAD51-4</i> mutants when treated with 0.025 $\mu\text{g}.\text{ml}^{-1}$ of phleomycin. ....	85
Table 3-5: Statistical analysis of the $\text{IC}_{50}$ values for methyl methanesulfonate (MMS) of the <i>RAD51-4</i> mutants.....	87
Table 3-6: Statistical analysis of the $\text{IC}_{50}$ values for phleomycin of the <i>RAD51-4</i> mutants.....	88
Table 3-7: Statistical analysis of the transformation efficiency of <i>RAD51-4</i> mutant cell lines. ....	89
Table 3-8: Percentage of <i>rad51-4</i> $\text{-/-}$ mutant cells with a specific number of RAD51 foci. ....	93
Table 3-9: Statistical analysis of the VSG switching frequencies of <i>RAD51-4</i> mutant cell lines. ....	99
Table 3-10: <i>in vitro</i> doubling times for <i>RAD51-6</i> mutant cell lines. ....	110
Table 3-11: Statistical analysis of the <i>in vitro</i> population doubling times for <i>RAD51-6</i> mutants. ....	110
Table 3-12: Statistical analysis of the clonal survival of <i>RAD51-6</i> mutants when treated with 0.0003 % methyl methanesulfonate (MMS).....	112
Table 3-13: Statistical analysis of the clonal survival for <i>RAD51-6</i> mutants when treated with 0.025 $\mu\text{g}.\text{ml}^{-1}$ phleomycin. ....	113
Table 3-14: Statistical analysis of the $\text{IC}_{50}$ values for MMS of the <i>RAD51-6</i> mutants. ....	115
Table 3-15: Statistical analysis of the $\text{IC}_{50}$ values for phleomycin of the <i>RAD51-6</i> mutants. ....	116
Table 3-16: Statistical analysis of the transformation efficiency of <i>RAD51-6</i> mutant cell lines. ....	118
Table 3-17: Percentage of <i>rad51-6</i> $\text{-/-}$ mutant cells with a specific number of RAD51 foci. ....	119
Table 3-18: Statistical analysis of the VSG switching frequency of <i>RAD51-6</i> mutant cell lines. ....	122
Table 4-1: <i>in vitro</i> population doubling times for the homozygous mutants of <i>RAD51</i> and the <i>RAD51</i> paralogues.....	130
Table 4-2: Statistical analysis of the <i>in vitro</i> population doubling times for the homozygous mutants of <i>RAD51</i> and the <i>RAD51</i> paralogues.....	131
Table 4-3: Statistical analysis of the $\text{IC}_{50}$ values of MMS for the homozygous mutants of <i>RAD51</i> and the <i>RAD51</i> paralogues.....	133
Table 4-4: Statistical analysis of the $\text{IC}_{50}$ values of phleomycin for the homozygous mutants of <i>RAD51</i> and the <i>RAD51</i> paralogues. ....	134
Table 4-5: Statistical analysis of the transformation efficiency for the homozygous mutants of <i>RAD51</i> and the <i>RAD51</i> paralogues.....	137
Table 4-6: Percentage of <i>RAD51</i> paralogue mutant cells with a specific number of RAD51 foci when treated with phleomycin. ....	140

Table 4-7: VSG switching frequencies of <i>rad51</i> and <i>RAD51</i> paralogue homozygous mutants.....	144
Table 7-1: The percentage of amino acid sequence identities and similarities when the <i>T. brucei</i> RAD51 paralogues are compared with <i>T. brucei</i> RAD51 and <i>H. sapiens</i> Rad51 and Rad51 paralogues.....	234
Table 7-2: The percentage of amino acid sequence identities when the <i>T. brucei</i> RAD51 paralogues were compared with the predicted <i>T. cruzi</i> and <i>L. major</i> RAD51 paralogues. ....	236
Table 7-3: RAD51 paralogue gene numbers found in eukaryotic organisms.....	237

## List of Figures

Figure 1-1: Light microscope images of <i>Trypanosoma brucei</i> among red blood cells. ....	2
Figure 1-2: Life Cycle of <i>Trypanosoma</i> species. ....	4
Figure 1-3: Variant Surface Glycoprotein (VSG). ....	10
Figure 1-4: Schematic representation of the structure of a bloodstream expression site (BES). ....	11
Figure 1-5: <i>in situ</i> transcriptional VSG switching event. ....	12
Figure 1-6: VSG switching occurs by recombination of different VSG genes from different locations on the <i>T. brucei</i> chromosome. ....	14
Figure 1-7: Mosaic VSG formation by segmental gene conversion. ....	15
Figure 1-8: Reciprocal exchange of the VSG gene resulting in a VSG switching event. ....	16
Figure 1-9: Pathways of DNA double strand break repair. ....	19
Figure 1-10: A simplified representation of the Non-Homologous End-Joining (NHEJ) pathway. ....	21
Figure 1-11: Model for the homologous recombination DNA repair pathway. ....	26
Figure 1-12: Sequence comparisons of prokaryotic RecA with eukaryotic Rad51 from yeast and humans. ....	33
Figure 1-13: Predicted model of complex formations and interactions among <i>S. cerevisiae</i> Rad51 and the Rad51 paralogues. ....	38
Figure 1-14: Predicted model of complex formations and interactions among mammalian Rad51 and the Rad51 paralogues. ....	40
Figure 3-1: <i>RAD51-4</i> gene disruption construct and gene disruption strategy. ....	70
Figure 3-2: Analysis of the <i>RAD51-4</i> gene disruption by PCR-amplification. ....	72
Figure 3-3: Analysis of <i>RAD51-4</i> gene disruption by Southern blot. ....	74
Figure 3-4: Analysis of the expression of the <i>RAD51-4</i> mRNA in the mutant cell lines by RT-PCR amplification. ....	75
Figure 3-5: Strategy for re-expression of the <i>RAD51-4</i> by recombination into the $\beta\alpha$ tubulin array. ....	76
Figure 3-6: Analysis of the expression of the <i>RAD51-4</i> mRNA in the <i>RAD51-4</i> re-expressor cell lines by RT-PCR amplification. ....	77
Figure 3-7: Analysis of the RAD51-4 protein levels in the <i>RAD51-4</i> re-expressor cell lines by western blot analysis. ....	78
Figure 3-8: Analysis of growth of <i>RAD51-4</i> mutants <i>in vitro</i> . ....	79
Figure 3-9: Replication and division of the nucleus and kinetoplast at specific cell cycle stages of <i>T. brucei</i> bloodstream form. ....	81
Figure 3-10: Clonal survival of <i>RAD51-4</i> mutants <i>in vitro</i> when treated with methyl methanesulfonate (MMS). ....	83
Figure 3-11: Clonal survival of <i>RAD51-4</i> mutants <i>in vitro</i> when treated with phleomycin. ....	84
Figure 3-12: IC <sub>50</sub> values for methyl methanesulfonate (MMS) for <i>RAD51-4</i> mutant cell lines. ....	86
Figure 3-13: IC <sub>50</sub> values for phleomycin in <i>RAD51-4</i> mutant cell lines. ....	87
Figure 3-14: Transformation efficiency assay for the <i>RAD51-4</i> mutant cell lines. ....	89
Figure 3-15: RAD51 sub-nuclear foci in <i>T. brucei</i> . ....	91
Figure 3-16: RAD51 sub-nuclear foci in <i>rad51-4</i> $\Delta/\Delta$ mutant cell lines. ....	92
Figure 3-17: Percentage of <i>RAD51-4</i> mutant cell lines with a specific number of the RAD51 foci when treated with 1.0 $\mu\text{g}.\text{ml}^{-1}$ of phleomycin. ....	94

Figure 3-18: Percentage of <i>RAD51-4</i> mutant cells with a specific number of the <i>RAD51</i> foci when treated with 2.0 $\mu\text{g}.\text{ml}^{-1}$ of phleomycin. ....	94
Figure 3-19: Use of the transgenic <i>T. brucei</i> strain 3174.2 allows characterisation of the <i>VSG</i> switching events.....	96
Figure 3-20: The mean <i>VSG</i> switching frequencies of the <i>RAD51-4</i> cell lines. ....	98
Figure 3-21: Analysis of the switching mechanisms of the <i>RAD51-4</i> mutant cell lines. ..	100
Figure 3-22: <i>RAD51-6</i> gene disruption construct and gene disruption strategy.....	101
Figure 3-23: Analysis of the <i>RAD51-6</i> gene disruption by PCR-amplification. ....	103
Figure 3-24: Analysis of <i>RAD51-6</i> gene disruption by Southern blot analysis.....	104
Figure 3-25: Analysis of the expression of the <i>RAD51-6</i> mRNA in the mutant cell lines by RT-PCR amplification.....	105
Figure 3-26: Strategy for re-expression of the <i>RAD51-6</i> by recombination into the $\beta\alpha$ tubulin array. ....	106
Figure 3-27: Analysis of the expression of the <i>RAD51-6</i> mRNA in the <i>RAD51-6</i> re-expressor cell lines by RT-PCR amplification.....	107
Figure 3-28: Analysis of the generation of the <i>RAD51-6</i> re-expressor cell lines by Southern blot.....	108
Figure 3-29: Analysis of growth of <i>RAD51-6</i> mutant cell lines <i>in vitro</i> . ....	109
Figure 3-30: Clonal survival of <i>RAD51-6</i> mutants <i>in vitro</i> when treated with methyl methanesulfonate (MMS). ....	111
Figure 3-31: Clonal survival of <i>RAD51-6</i> mutants <i>in vitro</i> when treated with phleomycin.....	113
Figure 3-32: $\text{IC}_{50}$ values of MMS for <i>RAD51-6</i> mutant cell lines.....	115
Figure 3-33: $\text{IC}_{50}$ values of phleomycin for <i>RAD51-6</i> cell lines .....	116
Figure 3-34: Transformation assay for the <i>RAD51-6</i> mutant cell lines. ....	117
Figure 3-35: <i>RAD51</i> sub-nuclear foci in <i>rad51-6</i> $-/-$ mutant. ....	119
Figure 3-36: Percentage of <i>rad51-6</i> $-/-$ mutant cells with a specific number of <i>RAD51</i> foci. ....	120
Figure 3-37: The mean of the <i>VSG</i> switching frequencies in <i>RAD51-6</i> mutant cell lines.....	121
Figure 3-38: Analysis of the switching mechanisms of the <i>RAD51-6</i> mutant cell lines. ..	123
Figure 4-1: Diagram showing a generic construct for the generation of <i>RAD51</i> and the <i>RAD51</i> paralogue mutants. ....	126
Figure 4-2: Diagram showing the knockout constructs for the generation of <i>rad51</i> and the <i>RAD51</i> paralogue mutant cell lines.....	128
Figure 4-3: Analysis of the growth of <i>RAD51</i> and <i>RAD51</i> paralogue mutants <i>in vitro</i> ....	130
Figure 4-4: $\text{IC}_{50}$ values of MMS for the homozygous mutants of <i>RAD51</i> and the <i>RAD51</i> paralogues. ....	133
Figure 4-5: $\text{IC}_{50}$ values of phleomycin for the homozygous mutants of <i>RAD51</i> and the <i>RAD51</i> paralogues. ....	134
Figure 4-6: Transformation assay for <i>RAD51</i> and the <i>RAD51</i> paralogue mutants.....	137
Figure 4-7: Detection of the <i>RAD51</i> foci by indirect immuno-fluorescence. ....	140
Figure 4-8: Percentage of the <i>RAD51</i> paralogue mutant cells with a specific number of <i>RAD51</i> foci.....	141
Figure 4-9: Western blot analysis showing <i>RAD51</i> expression in the <i>RAD51</i> paralogue mutant cell lines. ....	142
Figure 4-10: Compiled data for <i>VSG</i> switching frequencies of <i>rad51</i> and <i>RAD51</i> paralogue homozygous mutants.....	143
Figure 5-1: Model of the yeast two-hybrid system.....	150
Figure 5-2: Diagram showing the LexA fusions with <i>RAD51</i> and <i>RAD51</i> paralogues. ..	151

Figure 5-3: Diagram showing the V5-NLS-B42 fusions with RAD51 and the RAD51 paralogues. ....	153
Figure 5-4: Western analysis of LexA-RAD51 expression with V5-NLS-B42-RAD51 and the RAD51 paralogues in <i>S. cerevisiae</i> L40 co-transformants. ....	157
Figure 5-5: Histidine auxotrophy analysis of LexA-RAD51 with V5-NLS-B42- RAD51 and the V5-NLS-B42-RAD51 paralogues.....	158
Figure 5-6: X-gal filter lift analysis of LexA-RAD51 with V5-NLS-B42-RAD51 and the V5-NLS-B42-RAD51 paralogues.....	159
Figure 5-7: Quantification of $\beta$ -galactosidase activity of LexA-RAD51 with V5-NLS- B42-RAD51 and the V5-NLS-B42-RAD51 paralogues.....	160
Figure 5-8: Western analysis of LexA-RAD51-3 expression with V5-NLS-B42- RAD51 and RAD51 paralogues in <i>S. cerevisiae</i> L40 co-transformants.....	161
Figure 5-9: Histidine auxotrophy analysis of LexA-RAD51-3 with V5-NLS-B42- RAD51 and the V5-NLS-B42-RAD51 paralogues.....	162
Figure 5-10: X-gal filter lift analysis of LexA-RAD51-3 with V5-NLS-B42-RAD51 and the V5-NLS-B42-RAD51 paralogues. ....	163
Figure 5-11: Quantification of $\beta$ -galactosidase activity of LexA-RAD51-3 with V5- NLS-B42-RAD51 and the V5-NLS-B42-RAD51 paralogues.....	164
Figure 5-12: Quantification of $\beta$ -galactosidase activity of LexA-RAD51-4 co- expressed with pYesTrp2, V5-NLS-B42-RAD51 and each V5-NLS-B42-RAD51- paralogue fusion.....	166
Figure 5-13: A comparison of the RAD51 paralogue proteins highlighting the N and C terminal amino acid sequence domain used for truncation of RAD51-4.....	168
Figure 5-14: Diagram showing the LexA fusions with the RAD51-4 truncations. ....	169
Figure 5-15: Auto-induction of the LexA-RAD51-4 C and N termini truncated fusions co-expressed with pYesTrp2 (EV) assayed using the yeast $\beta$ -galactosidase assay kit.....	170
Figure 5-16: Auto-induction of the truncated LexA-RAD51-4 C and N termini truncated fusions co-expressed with pYesTrp2 (EV) assayed by histidine auxotrophy....	170
Figure 5-17: Western analysis of LexA-RAD51-5 expression with V5-NLS-B42- RAD51 and RAD51 paralogues in <i>S. cerevisiae</i> L40 co-transformants.....	171
Figure 5-18: Histidine auxotrophy analysis of LexA-RAD51-5 with V5-NLS-B42- RAD51 and the V5-NLS-B42-RAD51 paralogues.....	172
Figure 5-19: X-gal filter lift analysis of LexA-RAD51-5 with V5-NLS-B42-RAD51 and the V5-NLS-B42-RAD51 paralogues. ....	172
Figure 5-20: Quantification of $\beta$ -galactosidase activity of LexA-RAD51-5 with V5- NLS-B42-RAD51 and the V5-NLS-B42-RAD51 paralogues.....	173
Figure 5-21: Western analysis of LexA-RAD51-6 expression with V5-NLS-B42- RAD51 and RAD51 paralogues in <i>S. cerevisiae</i> L40 co-transformants.....	174
Figure 5-22: Histidine auxotrophy analysis of LexA-RAD51-6 with V5-NLS-B42- RAD51 and the V5-NLS-B42-RAD51 paralogues.....	175
Figure 5-23: X-gal filter lift analysis of LexA-RAD51-6 with V5-NLS-B42-RAD51 and the V5-NLS-B42-RAD51 paralogues. ....	175
Figure 5-24: Quantification of $\beta$ -galactosidase activity of LexA-RAD51-6 with V5- NLS-B42-RAD51 and the V5-NLS-B42-RAD51 paralogues.....	176
Figure 5-25: C-terminal HIS-tagged and N-terminal GST-tagged recombinant RAD51 and the four RAD51 paralogues proteins.....	177
Figure 5-26: Western analysis showing the specificity of the anti-RAD51 anti-serum. ...	179
Figure 5-27: Western analysis showing the specificity of the unpurified and affinity purified anti-RAD51 anti-serum. ....	181



Figure 5-28: Western analysis showing the specificity of the anti-RAD51-3 anti-serum.....	182
Figure 5-29: Western analysis showing the specificity of the affinity purified anti-RAD51-3 anti-serum.....	183
Figure 5-30: Western analysis showing the specificity of the anti-RAD51-4 anti-serum.....	184
Figure 5-31: Western analysis showing the specificity of the affinity purified anti-RAD51-4 anti-serum.....	186
Figure 5-32: Western analysis showing the specificity of the purified anti-RAD51-4 anti-serum using the acetone extraction method of total proteins from cells not expressing RAD51-4.....	187
Figure 5-33: Western analysis showing the specificity of the anti-RAD51-5 antibodies purified from chicken eggs. ....	188
Figure 5-34: Western analysis showing the specificity of the affinity purified anti-RAD51-5 anti-serum.....	189
Figure 5-35: Western analysis showing the specificity of the anti-RAD51-6 anti-serum.....	191
Figure 5-36: Western analysis showing the specificity of the affinity purified anti-RAD51-6 anti-serum.....	192
Figure 5-37: Western analysis of the eluates following the co-immunoprecipitation using affinity purified anti-RAD51-3 anti-serum. ....	194
Figure 5-38: Western analysis following the co-immunoprecipitation using affinity purified anti-RAD51-3 anti-serum.....	195
Figure 5-39: Western analysis of the eluates following the co-immunoprecipitation using affinity purified anti-RAD51-4 anti-serum. ....	196
Figure 5-40: Western analysis following the co-immunoprecipitation using affinity purified anti-RAD51-4 anti-serum.....	197
Figure 5-41: Schematic diagram showing the TAP-tag and PTP-tag.....	198
Figure 5-42: Overview of the PTP-purification strategy. ....	199
Figure 5-43: Diagram showing design and use of the C-terminal PTP vector for RAD51-5 tagging at its endogeneous locus.....	200
Figure 5-44: PCR-amplification to check PTP-tag integration into the <i>RAD51-5</i> allele...201	
Figure 5-45: Western analysis to confirm the integration of the PTP-tag into the endogenous RAD51-5 locus. ....	202
Figure 5-46: Diagram showing design and use of an N-terminal PTP vector for <i>RAD51-6</i> tagging at its endogeneous locus. ....	203
Figure 5-47: PCR-amplification to check PTP-tag integration into the <i>RAD51-6</i> allele...204	
Figure 5-48: Western analysis showing the PTP-tagging of RAD51-6 by probing the cell lysate with anti-PAP antibody.....	205
Figure 5-49: Coomassie stain and western blot of the RAD51-5-PTP purification. ....	207
Figure 5-50: Western analysis showing possible co-expression of RAD51-3 and RAD51-4.....	208
Figure 6-1: Transformation efficiency assay .....	213
Figure 6-2: Intra-chromosomal homologous recombination reporter constructs with direct and indirect repeats. ....	215
Figure 6-3: The positive control construct allowing luciferase expression from the $\beta\alpha$ -tubulin array. ....	217
Figure 6-4: Bioluminescence production in <i>T. brucei</i> bloodstream stage cells.....	218
Figure 6-5: Generation of a direct repeat luciferase construct to assay homologous recombination. ....	221

Figure 6-6: Integration of the luciferase construct into the $\beta\alpha$ tubulin array.....	222
Figure 6-7: Analysing recombination of the luciferase homologous sequence by PCR-amplification. ....	224
Figure 6-8: Analysing recombination of the luciferase homologous sequence by PCR-amplification after treatment with 0.0004 % MMS. ....	225
Figure 6-9: Luciferase construct and integration into the $\beta\alpha$ tubulin array. ....	227
Figure 7-1: An alignment of <i>T. brucei</i> RAD51, DMC1 and the RAD51 paralogues. ....	233
Figure 7-2: Synteny diagram between the orthologues of the surrounding genes of <i>T. brucei</i> RAD51-5 and <i>L. major</i> hypothetical protein. ....	238
Figure 7-3: An alignment of <i>T. brucei</i> and <i>T. cruzi</i> RAD51-5 with the predicted “RAD51-5” amino acid sequence of <i>L. major</i> , <i>L. braziliensis</i> , and <i>L. infantum</i> . ....	240
Figure 7-4: A phylogenetic tree of the Rad51 related proteins. ....	242
Figure 7-5: Models of the <i>T. brucei</i> RAD51 paralogue interactions. ....	253

“Darwin has shown us that we are not apart from the natural world,  
we do not have dominion over it, we are subject to its laws and processes  
as are all other animals on earth, to which indeed we are related.”

David Attenborough,

Charles Darwin and the Tree of Life, BBC.

## **Acknowledgement**

I wish to dedicate this work to my family, my Mum and Dad, who have always been there for me through thick and thin, and Barry, without his companionship and encouragement I would never have succeeded.

I would like to thank three important people who have helped to complete this work. Firstly, the guidance and constant support of Richard has made working at WCMP a pleasure. Secondly, I owe a great deal of thanks to Chris, who has added to my experimental and theoretical knowledge. Lastly, whenever an administrative problem arose, Alex was the person to fix it.

I would also like to thank my colleagues in the lab, past and present members of WCMP, many of whom I have become close friends, for the banter, gym and pub sessions.

I thank the Medical Research Council for funding and the Roberts Fund for a travel scholarship.

**Author's Declaration**

I declare that this thesis and the results presented within are entirely my own work except where otherwise stated. No part of this thesis has been previously submitted for a degree at any university.

---

Rachel Dobson

## **Abbreviations**

ATP	adenosine triphosphate
BER	base excision repair
BES	bloodstream expression site
BIR	break-induced replication
BLE	bleomycin resistance protein gene
bp	base-pairs
BSA	bovine serum albumin
BSF	bloodstream form
BSD	blasticidin-S-deaminase gene
cDNA	complementary DNA
co-IP	co-immunoprecipitation
DAPI	4', 6-diamidino-2-phenylindole
DIC	differential interference contrast
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
DSB	double strand break
dsDNA	double-stranded DNA
DTT	1, 4-dithiothreitol

EDTA	ethylenediaminetetraacetic acid
EM	electron micrograph
ESAG	expression site associated gene
ESB	expression site body
FITC	fluorescein isothiocyanate
GPI	glycosylphosphatidylinositol
HJ	Holliday junction
H-h-H	Helix-hairpin-Helix
HR	homologous recombination
HYG	hygromycin phosphotransferase gene
IR	intergenic region
kb	kilobase-pairs
kDa	kilo-dalton
LB	Luria-Bertani
Mb	megabase-pairs
MMEJ	microhomology mediated end-joining
MMS	methyl methanesulphonate
mRNA	messenger RNA

NEO	neomycin phosphotransferase gene
NER	nucleotide excision repair
NHEJ	non-homologous end-joining
ORF	open reading frame
PARP	procyclic acidic repetitive protein
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PM	polymerisation motif
PUR	puromycin-N-acetyltransferase gene
RNA	ribonucleic acid
RNAi	RNA interference
RNA pol I	RNA polymerase I
RNA pol II	RNA polymerase II
RPA	replication protein A
RT	reverse transcriptase
RT-PCR	reverse transcription polymerase chain reaction
SCC	sodium citrate with sodium chloride
SDS	sodium dodecyl sulphate



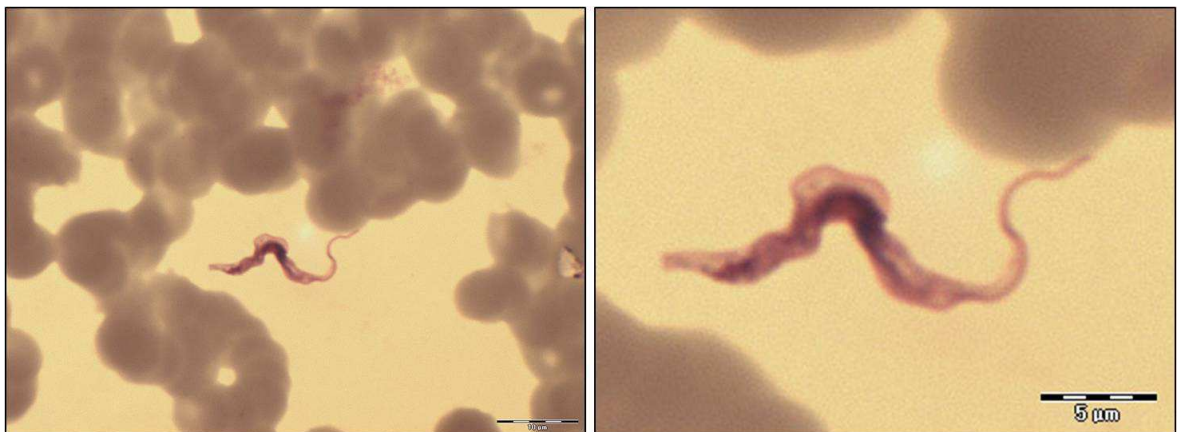
SDSA	synthesis-dependent strand annealing
SSA	single-strand annealing
ssDNA	single-stranded DNA
TAE	Tris-acetate-EDTA
TE	Tris-EDTA
TREU	Trypanosomiasis Research Edinburgh University
UV	ultraviolet
VAT	variable antigen type
VSG	variant surface glycoprotein
X-Gal	5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactoside

# **Chapter 1. Introduction**

## 1.1 General introduction to *Trypanosoma brucei*

### 1.1.1 Phylogeny of the African trypanosome

The genus *Trypanosoma* was first observed in 1841 by Gabriel Valentin; whilst examining blood from an infected brown trout, he detected the motile unicellular organism (Schmidt & Roberts, 1989). A century and a half of investigation of these microorganisms has defined *Trypanosoma* as members of the large and diverse protozoan kingdom. *Trypanosoma* are among the most ancient eukaryotic organisms known, diverging from *Homo sapiens* around 1.98 billion years ago (Hedges, 2002; McKean, 2003; Schmidt & Roberts, 1989). Taxonomically, they are ordered as Kinetoplastida, as their mitochondrial DNA is organised into a disc-like structure, the kinetoplast, located near the basal body (Maslov *et al.*, 2001; McKean, 2003).



**Figure 1-1: Light microscope images of *Trypanosoma brucei* among red blood cells.**

A *T. brucei* cell in a blood smear, stained using the Giemsa method. The image on the left shows the cell surrounded by red blood cells and the image on the right shows the cell in more detail, clearly showing the undulating membrane. Reproduced with the permission of B. Neish.

Within this order, the family *Trypanosomatidae* has nine genera (McGhee & Cosgrove, 1980), which parasitise vertebrates, invertebrates and plants, of which the so-called “TriTryps” are members, *Trypanosoma brucei*, *T. cruzi* and *Leishmania major*. *Trypanosoma* has two species which are human infective, the *T. cruzi* of South America, and *T. brucei* of Africa, while *Leishmania* species are endemic to South American and Asian countries (WHO, 2009). The majority of *Trypanosoma* species are heteroxenous, in that one stage of their life cycle is in the blood (see Figure 1-1) or tissue of a vertebrate and the other stage is in the intestine of a bloodsucking invertebrate (Schmidt & Roberts, 1989). The genus *Trypanosoma* is also grouped into two broad categories: the Salivaria and the Stercoraria, based on the characteristics of its development in the invertebrate host (Schmidt & Roberts, 1989). The three main Salivarian *Trypanosoma* species found in

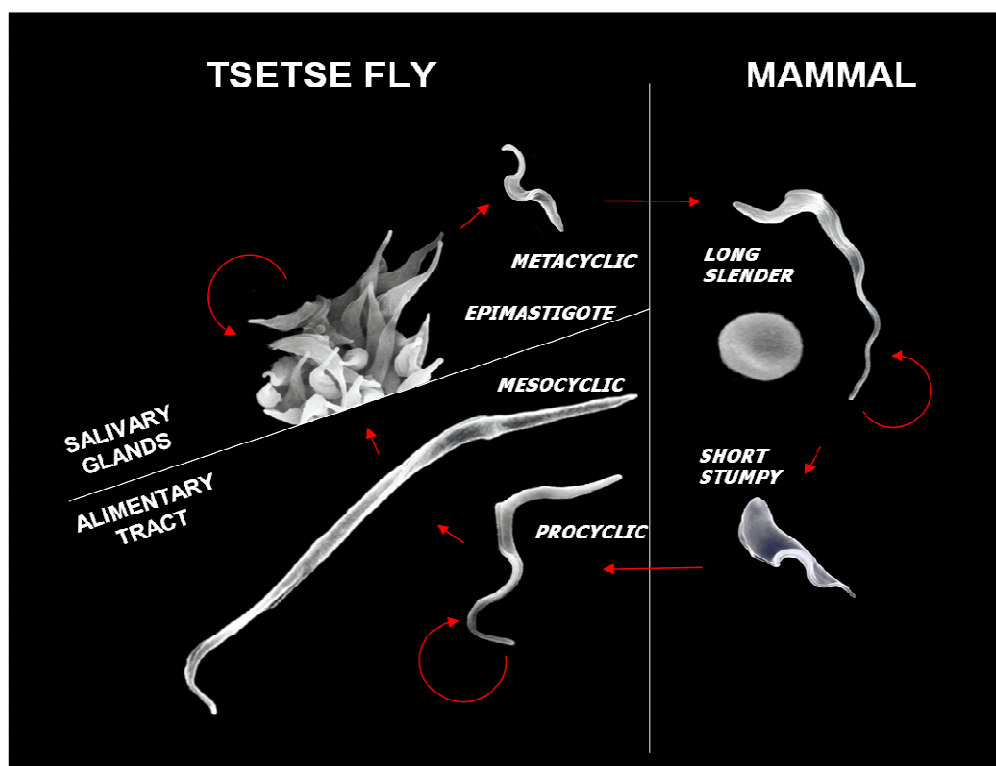
Africa are *Trypanosoma brucei*, *Trypanosoma congolense* and *Trypanosoma vivax* (Schmidt & Roberts, 1989; WHO, 2009). Of these three species, humans act as host to only one of them, *T. brucei*, whilst the others, *T. congolense* and *T. vivax*, are animal infective (Schmidt & Roberts, 1989).

### 1.1.2 Epidemiology and Life cycle

The species of *T. brucei* is pathogenically and geographically classified into three morphologically identical sub-species, *T. brucei*, *T. b. gambiense* and *T. b. rhodesiense* (Schmidt & Roberts, 1989). *T. b. brucei* is non-human infective, at least in part because it can be lysed by apolipoprotein L1, a trypanolytic factor in human serum (Pays *et al.*, 2006), whereas *T. b. gambiense* and *T. b. rhodesiense* have developed resistance to human trypanolysis (Turner *et al.*, 2004). *T. b. gambiense* is found in the riverine areas of west and central Africa while *T. b. rhodesiense* is found in the savannah of the eastern part of the continent (Hide *et al.*, 1990; Hide, 1999).

*T. brucei* is a haemoflagellate and proliferates extracellularly in mammals (Schmidt & Roberts, 1989; see Figure 1-1). It was realised, by David Bruce, that the parasite was insect-borne and infected members of the *Glossina* species, the so-called “Tsetse” fly (Cross, 2001; Steverding, 2008). Figure 1-2 shows the complex life cycle of the *Trypanosoma brucei* species (Barry & McCulloch, 2001). When an infective Tsetse fly has a blood meal from a vertebrate host, it deposits mature metacyclic cells in the connective tissue (Vickerman, 1985). In the mammalian host these parasitic cells undergo morphological changes while moving from the lymphatic system to the bloodstream and intracellular spaces (Matthews & Gull, 1994; Vickerman, 1985). From the non-dividing metacyclic form, the cells change into a long slender bloodstream form, which can proliferate by cell division (Matthews *et al.*, 2004). The population in the bloodstream is pleomorphic, in that the long slender form differentiates into an intermediate form and then two types of short stumpy bloodstream: one type which can develop as a procyclic form in the vector and the other that is programmed to die and are not tsetse infective (Seed & Wenck, 2003). The parasitaemia fluctuates partly due to the antigenic variation mechanisms of the parasite (reviewed in: Barry & McCulloch, 2001; see Section 1.2 below). Cells expressing a particular antigenic type form a part of the *T. brucei* population, called the homotype (Vickerman, 1985). An immune response mounted against this population results in death of this sub-set, while antigenic variant types survive. This causes waves of parasitaemia, where the majority of the cells are killed and the minority that switch live on (Morrison *et al.*, 2005).

Once differentiated into the bloodstream short stumpy form, the cells can adapt to life in the alimentary canal of the Tsetse fly. When ingested by the fly, the parasites move by undulation along the flagellum and, once in the mid-gut, form established procyclic cells. These cells then differentiate into proventricular mesocyclic forms, which move from the mid-gut to the salivary glands (Roditi & Liniger, 2002). Once in salivary glands, epimastigotes and prometacyclics form and attach by flagellipodium to the gland wall (Fenn & Matthews, 2007). When the cells detach from the gland wall and differentiate into metacyclics, the life cycle is complete and the parasite can be transmitted to another mammalian host via a second blood meal; see Figure 1-2.



**Figure 1-2: Life Cycle of *Trypanosoma* species.**

*T. brucei* life cycle stages are shown as scanning electron micrographs (EM), shown to scale; an erythrocyte is shown next to the long slender bloodstream stage. The host, Tsetse fly or mammal, and the name of the life cycle stage are labelled. Circular arrows represent replicative stages, whereas straight arrows represent differentiation and progression through the life cycle. Reproduced with permission of L. Tetley and D. Barry; adapted from Barry & McCulloch (2001).

### 1.1.3 Trypanosomiasis: symptoms & treatment.

Sleeping sickness, or Human African Trypanosomiasis (HAT), caused by *T. b. gambiense* and *T. b. rhodesiense*, is a fatal wasting disease if left untreated (Barrett *et al.*, 2003). Together these species are estimated to be responsible for more than 300,000 human cases per year (Sternberg, 2004; Stich *et al.*, 2002), with about 55 million at risk of infection (WHO, 2009). *T. congolense*, *T. vivax* and *T. b. brucei* are the cause of nagana in cattle, which has a great economic impact by destroying valuable livestock (WHO, 2009). *T. b.*

*gambiense* infection charts a chronically slow, but progressive course lasting anything up to several years. *T. b. rhodesiense* infection, in contrast, causes an acute form of the disease where progression to the later stages occurs within three months. Both infections have a similar clinical pathology and are fatal if left untreated. Once bitten by an infected Tsetse fly, the metacyclic parasites establish in the skin, differentiate in to the bloodstream stage and spread via the lymphatic system into the blood (Sternberg, 2004). The early haemolymphatic stage results in anaemia, weakness and weight-loss. In the late meningo-encephalitic stage, the parasite passes through the blood-brain barrier and infects the central nervous system (Enanga *et al.*, 2002), leading to sensory, motor and psychiatric disturbance, along with sleep abnormality (Lundkvist *et al.*, 2004).

Treatment efficacy is limited as there are few effective safe drugs available. There are four main drugs available for the management of the infection. Suramin sodium and pentamidine isethionate are used for the early stages before CNS involvement, and treat *T. b. rhodesiense* and *T. b. gambiense* infections respectively (Kennedy, 2004). In the advanced stages, the arsenical drug, melarsoprol is used for treatment of infections by both subspecies, while eflornithine (difluoromethylornithine; DFMO) is used for treatment of the disease caused by *T. b. gambiense* when the patients are unresponsive to melarsoprol (Kennedy, 2004). These drugs need to be administered intravenously or intramuscularly, and some cause a large group of side effects, including skin lesions, renal failure, cardiac abnormalities and fatal encephalopathy (Gull, 2002; Jannin & Cattand, 2004; Kennedy, 2004).

With these disadvantages and the fact that new drugs for treatment of trypanosomiasis are not being produced, the implementation of control strategies may be the best way of combating the disease in epidemic areas (Jannin & Cattand, 2004). The hope of development of an effective vaccine has somewhat been dismissed by the discovery of a complex antigenic variation system in trypanosomes. Nevertheless, understanding the mechanisms behind antigenic variation may lead to new drug targets as well as better knowledge of the recombination pathways of *T. brucei* and possibly of other eukaryotic cells.

#### **1.1.4 Use of *T. brucei* as a model organism?**

Scientists have used a small number of model organisms to attempt to simplify the study of complex molecular mechanisms (Hunter, 2008). A century of science has built-up on the belief that describing the sub-systems of a cell or organism will facilitate the understanding

of the whole system (this is reductionism). Another belief is that understanding the consequences of any two events will lead to a predictable result (this is causality). In the past decades, model organisms such as yeasts, rodents, *E. coli*, *Drosophila* and the nematode *C. elegans* have been used, representing only a small fraction of the biodiversity on the planet (Hedges, 2002). Extrapolation of findings from these experiments to all forms of life is potentially misleading or at least may not always be straightforward, yet a vast amount of useful data has been generated to describe the fundamental molecular mechanism supporting life such as replication, transcription, translation and DNA repair. Now, use of genome projects and development of molecular tools to examine other organisms, has allowed the metaphorical “research door” to open and more research opportunities to take place.

Research on *T. brucei* was initially carried out to add to our clinical knowledge of the parasite. Now, with the complete genome sequenced and our ability to culture and genetically manipulate the parasite, *T. brucei* can be regarded as a candidate for use as a model organism. As well as this, *T. brucei* can be used, in combination with a few others such as *Plasmodium falciparum*, *Toxoplasma gondii* and *Leishmania*, as model parasites. Also, it has been shown by Hedges *et al.*, 2004, that the sub-phylum of Euglenozoans (containing the TriTryps) diverged from other eukaryotes 1,840 +/-200 million years ago and are from an early branch of eukaryotes (Hedges *et al.*, 2004; Hedges, 2002). This knowledge may lead to insights regarding phylogeny, and most importantly, for this project, insights into gene retention and function.

### **1.1.5 Genome structure and gene expression by polycistronic transcription**

The “TriTryps” are members of the order Kinetoplastida, as they possess a kinetoplast for the mitochondrial DNA. The nuclear chromosomes are present as linear DNA molecules, unlike the content of the kinetoplast, which is organised as an interconnected network of circular DNA comprising mini-circles and maxi-circles (reviewed in: Lukes *et al.*, 2002). The nuclear genome of *T. brucei* contains 11 mega-base chromosomes that have a total size of ~ 26 Mb (0.9 to > 6 Mb individually) (Berriman *et al.*, 2005; Hertz-Fowler *et al.*, 2007). There are also intermediate-chromosomes and mini-chromosomes found in the nucleus of *T. brucei* (Berriman *et al.*, 2005). In most strains, there are 1-7 intermediate chromosomes of 200-700 kb and numerous (~100) mini-chromosomes of 30-150 kb (Alsford & Ersfeld, 2003; Melville *et al.*, 2000; Wickstead *et al.*, 2004). In 2005, the project to sequence the mega-chromosome DNA of *T. brucei* was completed (Berriman *et*

*al.*, 2005) using the genome of the *T. b. brucei* strain TREU (Trypanosomiasis Research Edinburgh University) strain 927/4 that expresses a single variable antigen type (VAT), derivative GUTat 10.1. Systemic sequencing of the intermediate- and mini-chromosomes has not yet been reported.

It has been shown that the *T. brucei* mega-base chromosomes contain the house-keeping genes structured in directional gene clusters (see below), including many tandemly repeated gene families such as tubulin, calmodulin and procyclic acidic repetitive proteins (PARP) (Berriman *et al.*, 2005; Boothroyd & Konuniecki, 1995). Interestingly, the majority of the sub-telomeric genes found on the mega-chromosomes are members of the Variant Surface Glycoprotein (VSG) gene family which contains >1000 members, whilst the intermediate chromosomes house the expression sites for bloodstream VSGs and the mini-chromosomes contain a diverse array of silent VSGs (Berriman *et al.*, 2005; Marcello & Barry, 2007a; Marcello & Barry, 2007b; Wickstead *et al.*, 2004). Approximately 9,068 protein coding genes have been annotated on the sequenced mega-chromosomes of strain TREU 927/4, and of these predicted genes only 50.5 % have been assigned a function through sequence homology (Berriman *et al.*, 2005).

Most of the genes of *T. brucei* are transcribed by RNA polymerase II (RNA pol II), with the exception of the VSG, rRNA and PARP genes which are transcribed by RNA polymerase I (RNA pol I) (Boothroyd & Konuniecki, 1995; Navarro & Gull, 2001). Unlike higher eukaryotes, in which protein coding genes undergo mono-cistronic transcription, *T. brucei*, and the other TriTryps (as well as some other lower eukaryotes, such as nematodes and trematodes) undergo poly-cistronic transcription as inferred from the convergence of the genes as directional gene clusters (Blumenthal, 2005; Boothroyd & Konuniecki, 1995; Evans & Blumenthal, 2000; Johnson *et al.*, 1987; Liang *et al.*, 2003). Long poly-cistronic transcribed units (pre-mRNA) contain information for many proteins which are separated by short spacer regions. Post-transcriptional modification is required to process these long primary units into individual mature mRNAs and occurs by *trans*-splicing and polyadenylation (Clayton, 2002). *Trans*-splicing is a process where two separate transcripts (pre-mRNA) are joined together and occurs with addition of a 5' cap called a spliced leader RNA at the 5' end of the pre-mRNAs (Sather & Agabian, 1985; Ullu *et al.*, 1993). The addition of a polyadenylated (poly(A)) tail at the 3' end of the pre-mRNA also occurs (Matthews *et al.*, 1994). These modified and mature mRNAs are then translated to amino acids.



## 1.2 Antigenic Variation

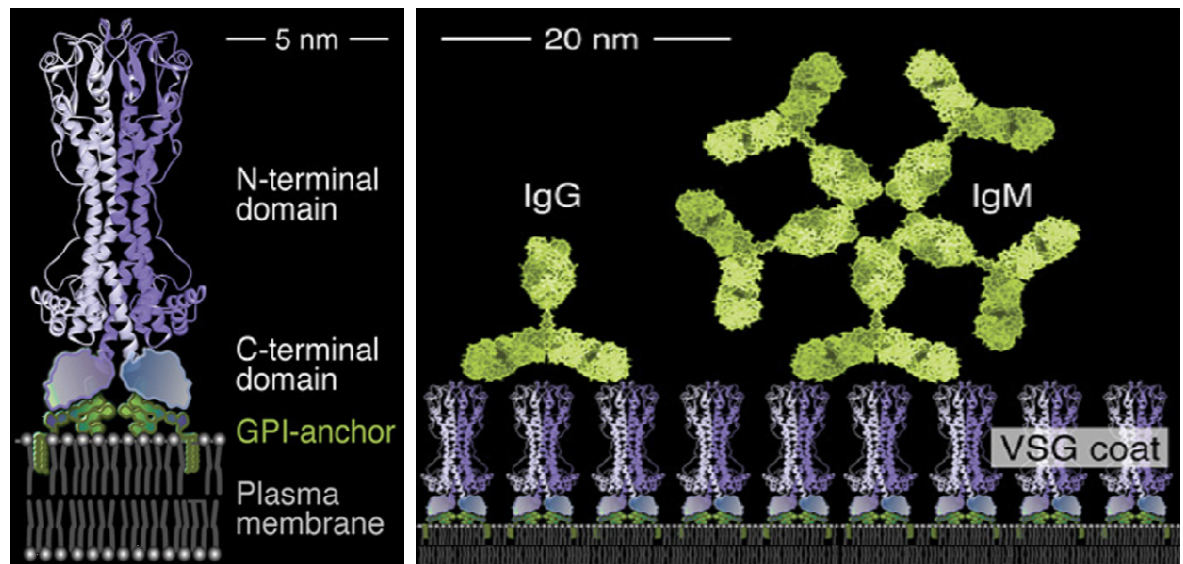
The success of most pathogens can, in part, be attributed to the organism's ability to overcome the protective immunity of the host. Scientists have uncovered a diverse collection of mechanisms used by bacteria, viruses and parasites to evade the onslaught of the host immune response. Some of these mechanisms include varying the antigens presented to the host immune system and/or modification of host immune responsiveness (Mims *et al.*, 1995: Chapter 2). Some bacteria find a safe haven from the immune response by invading the host cells, such as members of the genera *Listeria* and *Shigella* that replicate within the cell cytosol, whereas some *Legionella* and *Mycobacterium* species replicate within vacuoles of the cell (reviewed in: Wilson *et al.*, 2002). Parasites, such as *Plasmodium*, *Leishmania* and *T. cruzi*, also employ a similar tactic of evasion (Sibley, 2004). Microorganisms that do not invade the host cell, such as the extracellular *T. brucei*, are constantly surrounded by host immune factors and initially may seem more vulnerable than their intracellular relatives. This is not the case, however, as many microorganisms use phase or antigenic variation to change the antigens exposed on their surface, thereby altering their appearance to the host immune system (Barbour & Restrepo, 2000; Zambrano-Villa *et al.*, 2002). Phase variation refers to the reversible switch between two phenotypes, while antigenic variation refers to the irreversible, sequential expression of multiple different forms of the same antigen (van der Woude & Baumler, 2004).

Antigenic variation is observed widely in viruses, bacteria, fungi and parasites, with divergent underlying mechanisms (Deitsch *et al.*, 1997; Palmer & Brayton, 2007). Bacteria of the genus *Borrelia* use the mechanism of antigenic variation to change their exposed lipoprotein surface antigens termed variable large protein (*vlp*) and variable small protein (*vsp*) (Barbour *et al.*, 2006). A pattern of waves of spirochete infection occurs in the infected host, where a relapsing fever is characterised by high temperature and high bacteraemia, followed by a period of clearance (Norris, 2006). This cyclical pattern occurs when the majority of the bacterial population expressing a particular serotype (*vlp* or *vsp*) is cleared by specific host antibodies to that particular serotype (Norris, 2006). A new wave of bacteraemia results when a minority of the spirochetes have switched to an immunologically distinct lipoprotein surface antigen. The main mechanism of lipoprotein switching in *B. hermsii* and *B. burgdorferi* has been shown to be gene conversion, involving homologous recombination (HR) (Dai *et al.*, 2006; Zhang *et al.*, 1997). It has been found that in *B. hermsii* some 60 silent pseudogenes are present episomally, and gene conversion can occur between these fragments of *vlp* and *vsp* genes (Dai *et al.*, 2006).

*T. brucei* uses what appears to be a broadly related mechanism of antigenic variation to *Borrelia* (Palmer & Brayton, 2007). In *T. brucei*, the exposed antigen is a Variant Surface Glycoprotein (VSG) which covers the entire parasite surface and is switched to a different immunologically distinct antigen primarily by recombination, though transcriptional control mechanisms are also used, which is distinct from *Borrelia* (see Section 1.2.2; Barry & McCulloch, 2001). For *T. brucei* to evade the host's acquired immunity, the parasite must replace its VSG coat rapidly and completely with a new VSG that contains epitopes unrecognised by existing antibodies (Miller & Turner, 1981; Morrison *et al.*, 2005). When the antibodies eliminate a major part of the population, those parasites that have changed their VSG coat survive and proliferate. This pattern is repeated, resulting in fluctuating waves of chronic parasitaemia, with each wave containing *T. brucei* cells expressing distinct VSG coats (for reviews see Barry & McCulloch, 2001; Donelson, 2003; Pays, 2006). This process of VSG gene activation or replacement is sometimes called VSG switching, and it is this mechanism that allows a small number of parasites expressing the new VSG to survive the hostile extracellular environment of the mammalian host (Schmidt & Roberts, 1989). The mechanisms and potential regulation of VSG switching holds the key to the parasite's survival in the mammal, although the full picture of the factors involved seems to be elusive.

### **1.2.1 The role of Variant Surface Glycoproteins (VSG) in antigenic variation**

The dense surface coat covering the plasma membrane of pathogenic African trypanosomes was first imaged by high resolution micrographs (Vickerman, 1969). In the mammalian host, the surface coat of *T. brucei* is composed of  $5 \times 10^6$  dimers of Variant Surface Glycoprotein (VSG) (Auffret & Turner, 1981; Cross, 1975) and is attached to the plasma membrane by a glycosyl-phosphatidyl-inositol (GPI) anchor (Ferguson *et al.*, 1988).



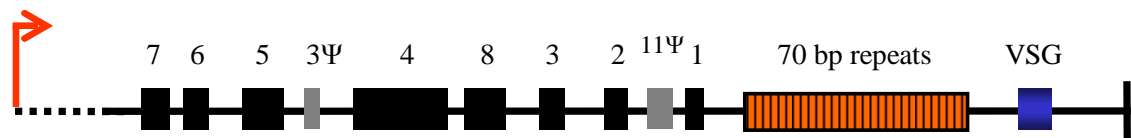
**Figure 1-3: Variant Surface Glycoprotein (VSG)**

The diagram on the left shows the 3D structure of a VSG dimer attached to a model of the *T. brucei* plasma membrane by a glycosyl-phosphatidyl-inositol (GPI) anchor. The diagram on the right models the densely packed VSG coat of the cell and predicted binding of IgG and IgM host antibodies. Reprinted with permission from Elsevier: Engstler *et al.* (2007), *Cell* **131**, pp 506.

It is this VSG coat which prevents the host immune response from targeting underlying invariant antigens (Borst & Fairlamb, 1998; Overath *et al.*, 1994). The VSG coat is immunogenic and is the point of attack for the host IgG and IgM responses (Engstler *et al.*, 2007; Turner *et al.*, 1988). Each VSG polypeptide within the mature VSG homodimer is between 400 to 500 amino acid residues in length, and is composed of a single N-terminal domain of 350-400 residues and one or two smaller C-terminal domains, each consisting of 40-80 residues (Carrington *et al.*, 1991; Chattopadhyay *et al.*, 2005; Pays, 2006; Pays *et al.*, 2007). The antibody responses are raised against the exposed hypervariable N-terminal domain of the VSG (Borst & Fairlamb, 1998; Field & Boothroyd, 1996). Although there is limited sequence conservation of the N-terminal domain, its structure is conserved and consists of two long  $\alpha$  helices, which interact with the neighbouring VSG in the homodimer (Blum *et al.*, 1993); see Figure 1-3. It appears that the buried C-terminal domain is more conserved, where the VSG is anchored to the plasma membrane and is less accessible to immune attack (Ferguson *et al.*, 1988).

The *T. brucei* genome has a size of ~26 Mb (see Section 1.1.5) and contains a large archive of over 1000 silent VSG genes, which are present in the sub-telomeric arrays of the chromosomes (Barry *et al.*, 2005). Some of these genes are also present in the VSG expression sites (ES) which are found in the megabase and intermediate chromosome telomeres (Berriman *et al.*, 2005; Hertz-Fowler *et al.*, 2007; Marcello & Barry, 2007a; Marcello & Barry, 2007b). Out of these silent VSG genes, 806 have been functional

annotated and only 7 % were found to encode fully functional VSGs, 9 % were atypical VSGs which are predicted not to be functional VSGs, 66 % were pseudogenes and 18 % were VSG gene fragments (Berriman *et al.*, 2005). The functionality of the VSGs found at the telomeres of the some 100 mini-chromosomes has not been determined (Berriman *et al.*, 2005; Pays *et al.*, 2007). In the VSG bloodstream stage expression sites, VSG pseudogenes are found, but the VSG adjacent to the telomere appears always to be functional (Barry *et al.*, 2005; Hertz-Fowler *et al.*, 2008). Although there is a large archive of silent VSG genes in the *T. brucei* genome, at any given time, only one single VSG gene is expressed from a VSG expression site (ES) (Barry & McCulloch, 2001). The ES is adjacent to the chromosomal telomere and the VSG gene is transcribed poly-cistronically with a series of other expression-site associated genes (*ESAGs*), which are normally separated from the VSG by a region of repeats each of 70 bp (Berriman *et al.*, 2002; Borst *et al.*, 1996) and are discussed below, in Section 1.2.2; Figure 1-4.

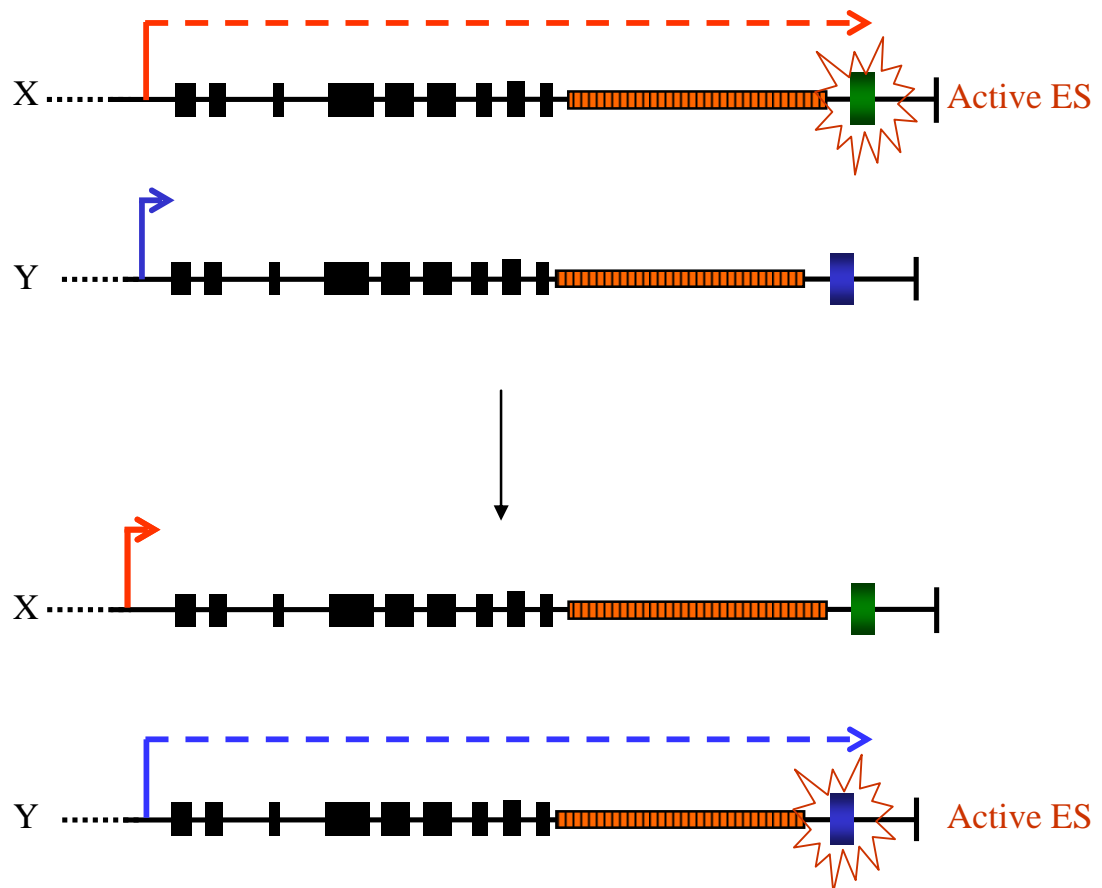


**Figure 1-4: Schematic representation of the structure of a bloodstream expression site (BES).**

The diagram shows an example of a *T. brucei* Lister 427 bloodstream expression site and indicates the variant surface glycoprotein (VSG; blue box) and expression-site associated gene (ESAG 1, 2, 3, 4, 5, 6, 7, 8, 11; black boxes) separated by the region of 70 bp repeats (orange boxes). Also shown are the promoter (red arrow), telomeric repeat region (vertical black line), and pseudo-*ESAGs* (Ψ). Adapted from Hertz-Fowler *et al.* (2008).

## 1.2.2 Mechanisms of VSG switching

The switching events that activate a previously non-expressed VSG gene occur independently of the host immune response (Doyle *et al.*, 1980) suggesting this is a pre-emptive survival strategy driven by the parasite. In the majority of cases, VSG switching occurs by gene replacement of the VSG in the ES with another VSG by DNA recombination (Hoeijmakers *et al.*, 1980; Robinson *et al.*, 1999). However, it has also been shown that there can be so-called *in situ* transcriptional switching events (Borst & Ulbert, 2001; Myler *et al.*, 1984), where one previously inactive ES is switched on to allow complete transcription of the VSG and *ESAGs* while the active ES is transcriptionally silenced (Figure 1-5).



**Figure 1-5: *in situ* transcriptional VSG switching event**

An *in situ* VSG switching event occurs when an active expression site (ES) is transcriptionally switched off, called X and a previously inactive ES is switched on, called Y. This results in the green VSG not being actively transcribed while the blue VSG is transcribed and consequently expressed along with the expression site associated genes: black boxes and 70 bp repeat region: orange boxes. The promoters are indicated by the red or blue arrows.

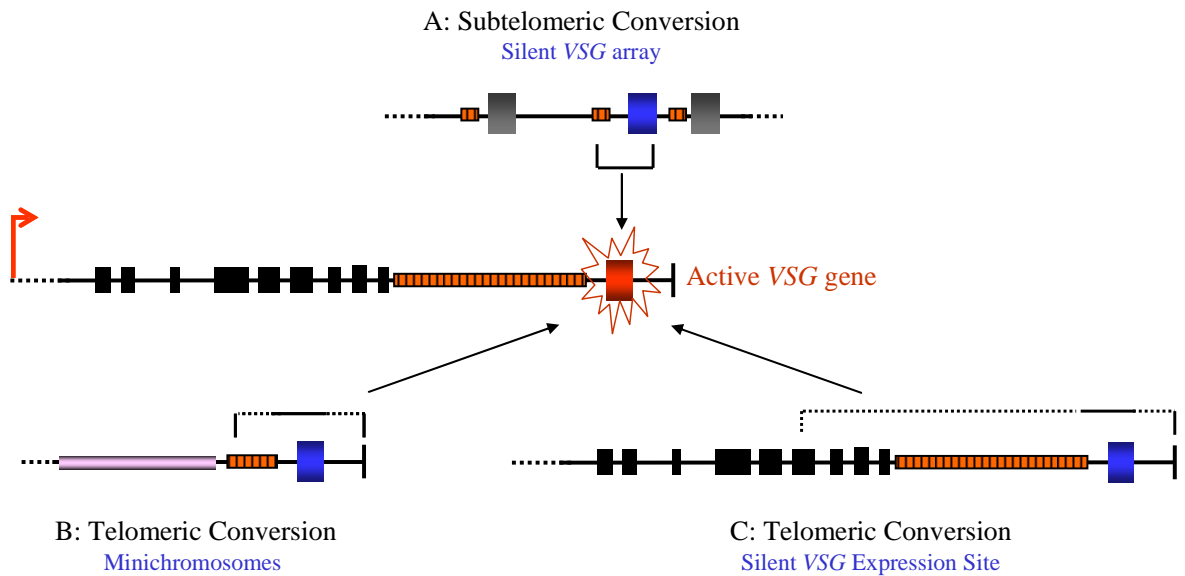
This can occur because there are multiple ESs within the *T. brucei* genome (Borst & Ulbert, 2001). Some of these are active for VSG expression in metacyclic cells (Rudenko *et al.*, 1994), while the bloodstream VSG expression sites (BES) are specifically transcribed in the mammal (Chaves *et al.*, 1999). A number of BES, from a few *T. brucei* isolates have now been sequenced, showing that their size, measured from the RNA pol I promoter to the telomere can span between 40-100 kb (Hertz-Fowler *et al.*, 2008; Young *et al.*, 2008). Within the BES there are at least eight expression site associated genes (ESAGs), only some of which have been assigned functions (Becker *et al.*, 2004; Berriman *et al.*, 2002). For example, two ESAGs (ESAG6 and ESAG7) encode the subunits for a transferrin receptor required for iron uptake (Schell *et al.*, 1991; Steverding *et al.*, 1995). Another ESAG, not found in all BES, encodes the serum resistance-associated (SRA) factor in *T. b. rhodesiense* (Van Xong *et al.*, 1998). The mechanisms dictating how one particular BES is fully transcribed while the others remain largely silent are unknown. A distinct sub-nuclear structure, named the expression site body (ESB), has been found to contain the actively transcribed VSG expression site along with RNA pol I, while the silent BESs are

found outside this ESB (McCulloch, 2004; Navarro & Gull, 2001). The discovery of the ESB may lead to an explanation for the controlling mechanisms for the expression of the active *VSG*. However, chromatin modification also appears to have a role, since mutation of the histone methyltransferase, DOT1B, causes de-repression of silenced BES and alters the dynamics of transcriptional switching (Figueiredo *et al.*, 2008; Stockdale *et al.*, 2008).

The most common mechanism for *VSG* switching involves DNA recombination events (Robinson *et al.*, 1999). There are three different recombination pathways of *VSG* switching, which are defined by the source of the donor DNA. All pathways rely on some sequence homology being present between the donor and recipient DNA.

The first *VSG* switching mechanism involving DNA recombination is duplicative transposition, or gene conversion. Gene conversion involves creating a new copy of a silent *VSG* and moving it into the BES, where it replaces the previously expressed *VSG*. This is the most frequently observed pathway experimentally (Robinson *et al.*, 1999). Silent *VSGs* in a number of genomic locations can act as gene donors (Figure 1-6; detailed below). The amount of *VSG* sequence copied during gene conversion is variable and normally extends beyond the boundaries of the *VSG* open reading frame. The upstream limit of gene conversion is often found to be 70 bp repeats (Liu *et al.*, 1983), that flank >90 % of *VSGs* (Marcello & Barry, 2007a) but can extend into the conserved *ESAGs*. The extent of the gene conversion event can vary downstream from the 3' end of the *VSG* (Michiels *et al.*, 1983), to the repeats of the telomere (De Lange & Borst, 1982).

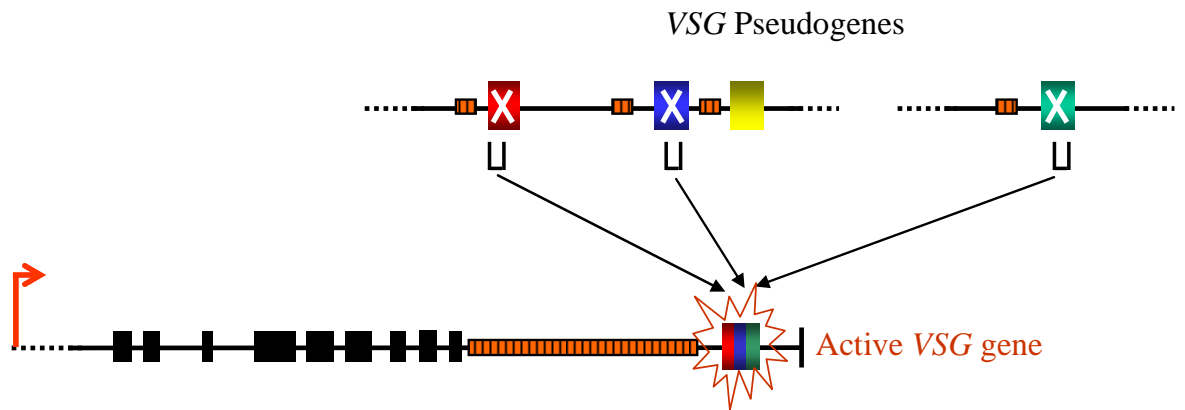
Gene conversion can occur by copying from the *VSG* sub-telomeric array of the mega-base chromosomes, where the 70 bp repeat region and 3' end of the *VSG* are used most commonly as recombination homology (see Figure 1-6A) (Liu *et al.*, 1983; Michiels *et al.*, 1983). Other gene donors for gene conversion include the telomeric *VSG* genes on either the mini-chromosomes or in the BES, either in the mega-base or intermediate chromosomes (see Figure 1-6B and 1-6C respectively). Though *VSG* gene conversion events have been shown to use the 70 bp repeat region upstream of the *VSG* gene as a substrate for the recombination event (the orange region in Figure 1-6) (De Lange *et al.*, 1983; Shah *et al.*, 1987), deletion of this region within a BES does not compromise the antigenic variation mechanism (McCulloch *et al.*, 1997) underlining the flexibility in substrate choice for this process.



**Figure 1-6: VSG switching occurs by recombination of different VSG genes from different locations on the *T. brucei* chromosome.**

Gene conversion copies a silent VSG (blue box) into an active expression site (red box) replacing the resident VSG. The amount of the sequence copied is depicted by the brackets. The dashed lines show variation in the amount of sequence that may be copied. Expression site associated genes: black boxes; 70 bp repeats: orange boxes; BES promoter: red arrow; and telomeric repeats: vertical black line.

The second mechanism of VSG switching is mosaic VSG gene formation. This is where segments of VSG pseudogenes are assembled most likely by segmental gene conversion reactions, to create a new functional VSG, that is subsequently expressed from the BES (as shown in Figure 1-7) (Roth *et al.*, 1986). It appears that this reaction is dependent on VSG ORF sequence homology and does not rely on flanking homology, though it may be that the reactions are anchored by the 70 bp repeat upstream or by downstream homology. Mosaic formation occurs later in infection when the chronically-infected host generates comprehensive immunity against most intact VSGs (Kamper & Barbet, 1992; Marcello & Barry, 2007a). However, since most of the VSG archive consists of pseudogenes (66 % of analysed VSGs are pseudogenes) or VSG gene fragments (18 %) it is suggested that the mosaic formation is important for the parasite's survival (Berriman *et al.*, 2005; Kamper & Barbet, 1992; Marcello & Barry, 2007a).

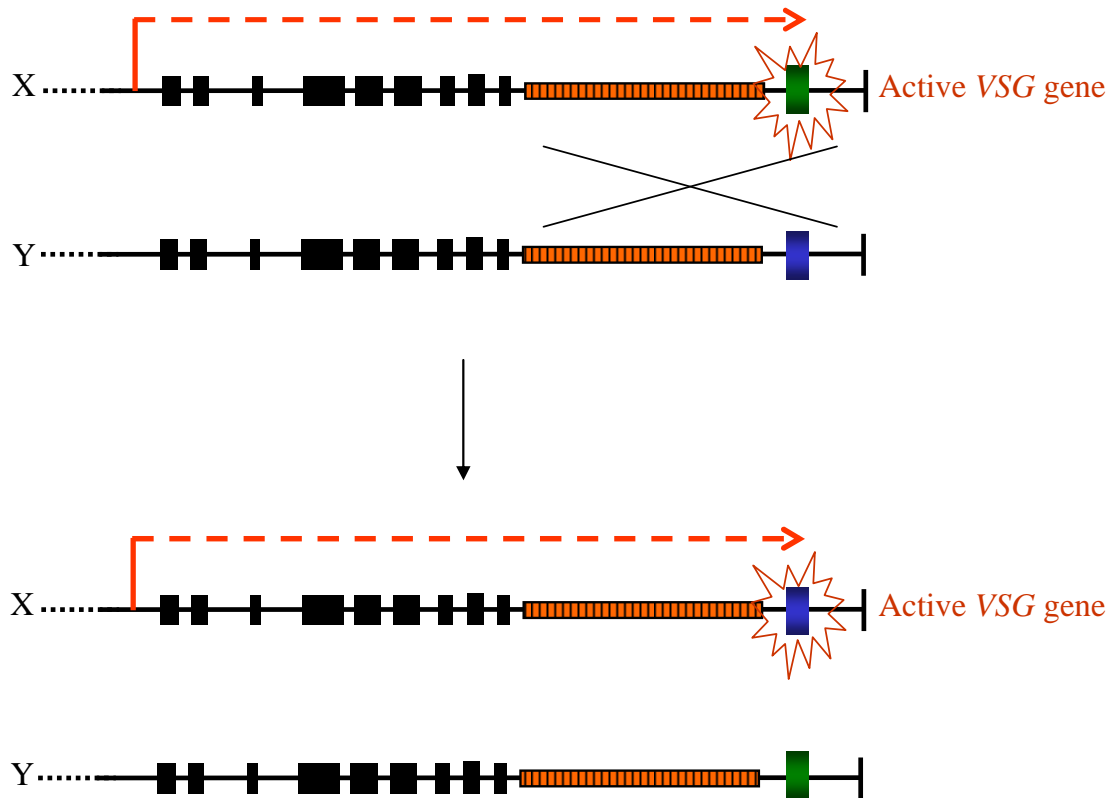


**Figure 1-7: Mosaic VSG formation by segmental gene conversion.**

Generating a new VSG from a number of silent VSG pseudogenes. The extent of the sequence copied is indicated by the brackets. Expression site associated genes: black boxes; 70 bp repeats: orange boxes; bloodstream expression site promoter: red arrow; and telomeric repeats: vertical black line.

The third mechanism of VSG switching occurs by reciprocal exchange. This is a classical HR event where a crossover occurs (Pays *et al.*, 1985; Shea *et al.*, 1986), switching a telomeric VSG gene (either from a silent BES or a mini-chromosome) into the active BES and moving the previously active BES into the other telomere. There is no loss or gain of sequence, as the chromosome ends containing the active expression site and a silent VSG are simply exchanged (see Figure 1-8). This mechanism is less common in VSG switching (Rudenko *et al.*, 1996), as it cannot act on the majority of the archive (the silent array in the mega-base chromosomes) as it is presumably limited to the telomeric VSGs since exchange of a VSG array by this approach is likely to be lethal (Borst *et al.*, 1996).





**Figure 1-8: Reciprocal exchange of the VSG gene resulting in a VSG switching event.**

Reciprocal exchange occurs where HR of the VSG gene and/or surrounding sequences can occur (crossed lines). The previously silent VSG gene (blue box) recombines into the active expression site (called X). There is no loss of DNA during this cross-over event and both VSG genes are retained. Expression site associated genes indicated by: black boxes; 70 bp repeats: orange boxes; expression site promoter: red arrow; and telomeric repeats: vertical black line.

These switching mechanisms enable the expression of new epitopes that have not been previously experienced by the host and allow the longevity of the parasite in the bloodstream. Significantly, accumulating evidence suggests that at least VSG gene conversion is driven by HR, as mutation of core factors such as RAD51, BRCA2 and RAD51-3, impair VSG switching (discussed below, Section 1.5).

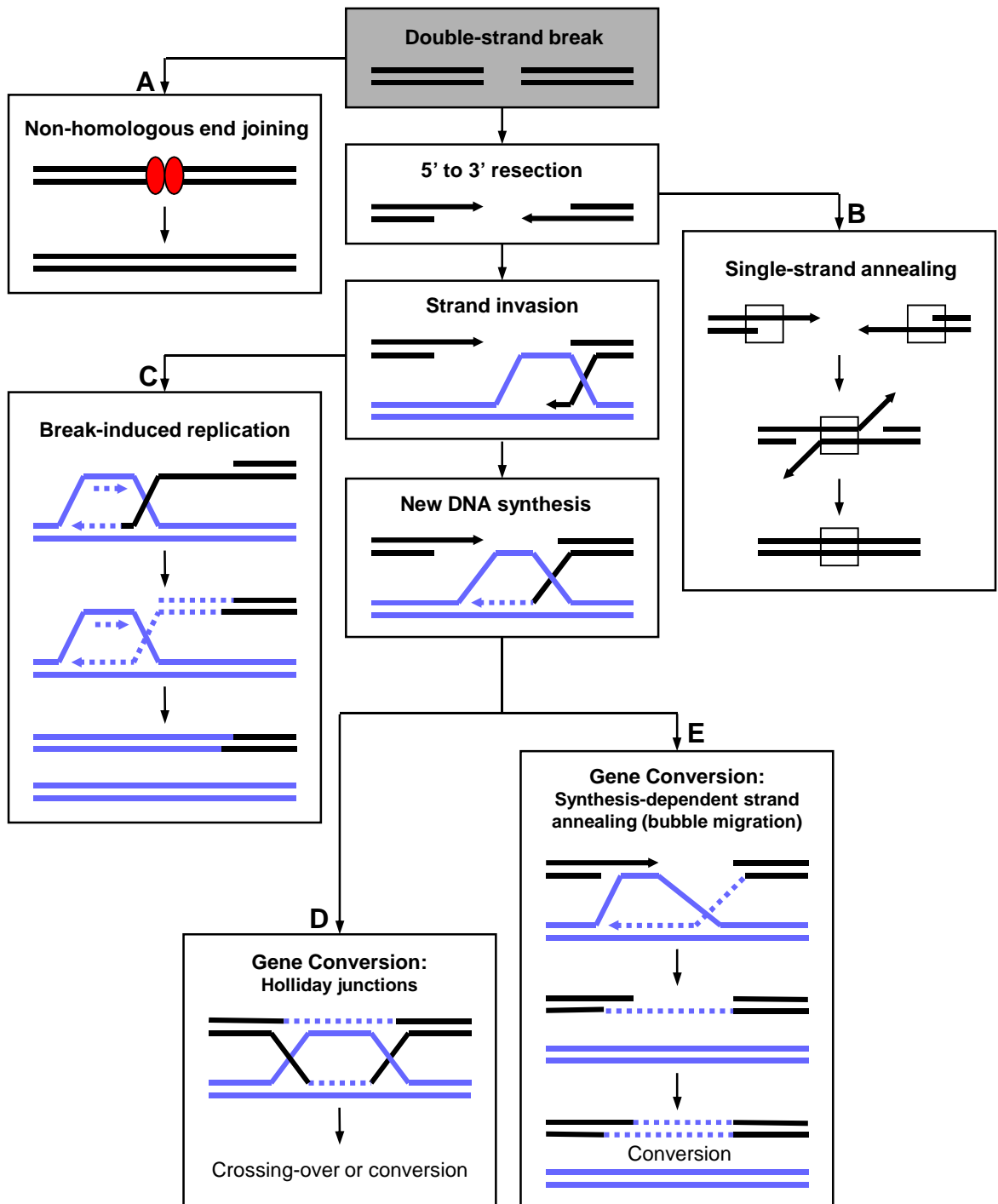
## 1.3 DNA repair

The genetic information encoded by genomes composed of nucleic acid molecules, DNA and RNA, allow information to be stored, transferred between generations and enable the production of RNA, proteins, and ultimately every cellular component (Lehninger *et al.*, 1993). As a result, the fidelity and preservation of the genome is extremely important. Endogenous and exogenous damaging agents, as well as the intrinsic instability of the DNA molecules themselves, result in a high level of damage to the DNA (Friedberg *et al.*, 1995). Endogenous damaging agents are mostly the by-products of metabolic pathways, such as oxidative metabolism, which results in hyperreactive oxygen molecules (hydrogen peroxide, peroxide and hydroxyl radicals), methylation and alkylation (Marnett & Plastaras, 2001; Pluskota-Karwatka, 2008). Exogenous damaging agents or environmental factors such as ionising radiation (X-rays and  $\gamma$ -rays), UV and chemical damaging agents can cause atypical bonding between the bases on the same strand, inter-strand crosslinks and breaks in the phosphodiester backbone. The intrinsic instability of the components of the DNA strand is also responsible for some DNA damage. Instability in the glycosyl-bond linking the pentose sugar ring to the base can cause the loss of that base and results in apurine or apyrimidine formation (Friedberg *et al.*, 1995). DNA interactions with proteins can also cause damage and instability; for example, errors in the replication of the new DNA strand can result in mismatches, insertions or deletions of bases. Stalling of the replication fork can also cause DNA damage (Chang & Cimprich, 2009). Finally, some processes in the cell deliberately create DNA damage, such as immunoglobulin and T cell receptor gene arrangements in mammals, and meiosis in many eukaryotes (discussed in Section 1.5.4 below).

Any DNA damage needs to be correctly repaired to maintain the fidelity of the genome. Repair of damaged DNA prevents chromosomal fragmentation or rearrangement, carcinogenesis in multicellular organisms and potentially cell death (Hakem, 2008). Consequently, all organisms have the ability to repair their DNA, which may be either error-free or error-prone. Given the diversity in the forms of DNA damage many different DNA repair pathways are used, which are conserved to a greater or lesser extent among all the phylogenetic groups (Eisen & Hanawalt, 1999). Broadly, these repair pathways can be divided into two groups; the first group uses excision mechanisms to repair the misplaced base(s) or distorted strand(s), and the second group repairs phosphodiester strand breaks, either single-stranded or double-stranded breaks.

Three main excision pathways of DNA repaired have been described; mismatch repair (MMR), base excision repair (BER), and nucleotide excision repair (NER) (review in: Christmann *et al.*, 2003; Kunz *et al.*, 2009; Robertson *et al.*, 2009). All of these have two common properties: the use of a nuclease to excise or remove the damaged base(s) and a DNA polymerase and ligase to fill the resulting gaps with the correct bases and re-seal the DNA. Although these two properties are common between the different mechanisms, the specific proteins used to carry out the repair in each pathway differ greatly. These pathways will not be discussed further here.

The second group of repair pathways, which act on double-strand DNA breaks has two primary distinct mechanisms (see Figure 1-9); non-homologous end-joining (NHEJ) and homologous recombination (HR) with the latter implicated in VSG switching, (see below Section 1.5 and above Section 1.2.2). As their name suggests, HR requires a region of homologous sequence to repair the DNA break while NHEJ does not (Symington, 2002). As HR requires base pairing with an unbroken template, normally the sequence integrity of the broken DNA molecule is maintained, whilst NHEJ involves re-ligation of the break ends and therefore, may result in loss of bases during the repair (see Section 1.4). HR can occur through different pathways depending on whether or not the DNA break ends invade a homologous DNA duplex, a process termed strand invasion. If the broken DNA ends do not invade the duplex DNA, but utilise base pairing around the DSB, this pathway is termed single-strand annealing (Figure 1-9B). If strand invasion occurs the HR pathway used depends on whether one or both strand ends invade the duplex DNA (Figure 1-9C, D, E) and these pathways will be discussed in more depth in Section 1.4 and Section 1.5 below.



**Figure 1-9: Pathways of DNA double strand break repair.**

A schematic representation of the two main mechanisms of double strand break (DSB) repair. The first is non-homologous end joining (NHEJ) (**A**) and the second is homologous recombination (HR) (**B-E**). DNA containing a DSB is represented by black lines, intact duplex DNA by blue lines, newly synthesised DNA by dashed lines, and NHEJ machinery by red circles. **B**: Single strand annealing (SSA) where the boxes indicate homology, **C**: Break induced replication (BIR), **D-E**: Gene conversion. Adapted from J.S. Bell, PhD thesis, 2002.

## 1.4 End-Joining Double Strand Break repair

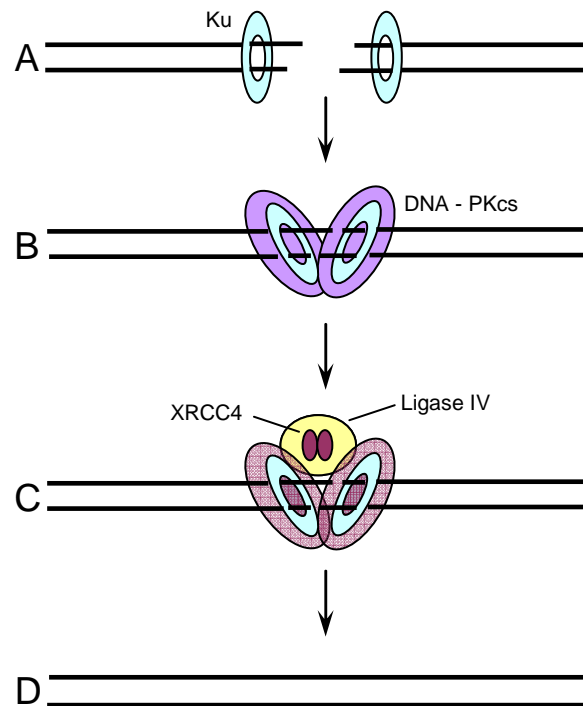
The repair of double strand breaks (DSB) can occur by re-joining the two ends of the broken DNA. To achieve this, there is the classical non-homology end joining (NHEJ) repair pathway, which may or may not require DNA homology (0-5 bp) to facilitate repair (Roth & Wilson, 1986). In addition, there is another putative end-joining repair pathway, which acts by joining the ends through the use of DNA micro-homology (5-25 bp) around the DSB (McVey & Lee, 2008).

### 1.4.1 Non-Homologous End Joining (NHEJ)

As discussed above (Section 1.3), the NHEJ pathway involves ligating the DSB ends together and frequently results in changes in the DNA sequence and may result in loss or addition of sequence around the break ends (Pfeiffer *et al.*, 1994; Roth & Wilson, 1986). Nevertheless, NHEJ has a major contribution in repair of the DSBs in mammalian cells, and is vital for cell survival (reviewed in: Dudasova *et al.*, 2004; Hefferin & Tomkinson, 2005; Weterings & van Gent, 2004).

The NHEJ pathway uses five primary protein components. Three are subunits of the DNA-dependent protein kinase, comprising the Ku70 and Ku80 proteins, which form a heterodimer DNA end-binding complex (Ku), and the catalytic subunit (DNA-PKcs) (Gottlieb & Jackson, 1993; Yaneva *et al.*, 1997). The other components act as a multiprotein complex comprising a NHEJ-specific ligase, DNA ligase IV and a ligase IV interacting factor, the XRCC4 protein (Critchlow *et al.*, 1997). Together these factors catalyze the DNA ligation step in the pathway. It has been shown that the Ku heterodimer binds with high affinity to both ends of the DSB, and once bound forms a bridge over the gap, leading to the recruitment of the other NHEJ factors (Bliss & Lane, 1997; Falzon *et al.*, 1993; Ramsden & Gellert, 1998) (see Figure 1-10A). The first to be recruited is DNA-PKcs (see Figure 1-10B). There is some evidence that DNA-PKcs phosphorylation alters or releases the DNA-PKcs DNA interaction (Block *et al.*, 2004; Kurimasa *et al.*, 1999). However, it is not yet fully understood the role that DNA-PKcs phosphorylation plays in NHEJ, nor whether DNA-PKcs autophosphorylates itself or whether it is phosphorylated by other means. DNA-PKcs is a member of the phosphatidylinositol 3-kinase-like kinase group (PIKK), of which ATM and ATR are also members (Weterings & Chen, 2007). Both ATM and ATR respond to DNA damage and may even act in the NHEJ pathway (Falck *et al.*, 2005). Binding of DNA-PKcs allows the recruitment of other NHEJ factors such as the Artemis nuclease (Goodarzi *et al.*, 2006). DNA Ligase IV-XRCC4 (see Figure

1-10C) is then recruited and the gaps in the DNA are processed (or filled-in) and ligated together, resulting in the break being repaired (see Figure 1-10D).



**Figure 1-10: A simplified representation of the Non-Homologous End-Joining (NHEJ) pathway.**

Repair of a DSB by the NHEJ pathway. **A:** A DSB is recognised by the Ku70/80 heterodimer (Ku). **B:** DNA-dependent protein kinase catalytic subunit (DNA-PKcs) is recruited. **C:** Processing of the ends follows with the LigaseIV-XRCC4 complex repairing the ends by ligation. **D:** The DSB is repaired. Other potential auxiliary factors in NHEJ such as the Artemis nuclease and Nej1/Lif1 in yeast and Cernunnos/XLF in mammals are not shown. Adapted from Weterings & Chen (2007).

It has been shown that NHEJ is the main pathway in V(D)J recombination (see review: Gellert, 2002; Lieber *et al.*, 2004). This recombination creates the diversity among the immunoglobulin (Ig) and the T cell receptor (TCR) genes found in the mammalian immune system (Abbas *et al.*, 1994: Chapter 4). V(D)J recombination is the process where the variable (V) gene segments are joined to the segments encoding the constant (or non-variable) regions of the Ig or TCR polypeptides. The variable gene segments can join to any number of the diversity (D) and joining (J) gene segments, and once joined, the three gene segments make up a new exon for generation of Ig and TCR heavy chains (Abbas *et al.*, 1994: Chapter 4). As an Ig or TCR locus may contain tens or hundreds of V, D, and J segments, the different combinational possibilities are great. As a result, there is a large number of possible Ig or TCR molecules and the combinational possibilities could be as high as  $10^7$ , allowing a diverse antibody response to many different agents (Gellert, 2002). There are specific recognition sequences between the gene segments called recombinational signal sequences (RSS). The RSS are moderately conserved and contain up to 7 or 9 conserved nucleotides separated by 12 or 23 non-conserved nucleotide spacers

(Tonegawa, 1983). The rearrangement occurs by excision of DNA sequence between the V, D, and J segments using the RSS as a guide, followed by ligation of the DNA ends (Dudley *et al.*, 2005). It has been found that there are V(D)J specific proteins for excision at specific recombinational sequences and looping of the gene segments, named RAG1 and RAG2 (recombination activation gene) (Oettinger *et al.*, 1990; Schatz *et al.*, 1989). These proteins generate DSBs, which recruits Ku, DNA-PKcs, and Ligase IV-Xrcc4, causing NHEJ to have a prominent role in the joining of the segment ends (Dai *et al.*, 2003).

### 1.4.2 Micro-homology Mediated End-Joining (MMEJ)

An accumulating number of reports have shown a distinct form of DSB end-joining mediated by short stretches of homology (5-25 bp) (reviewed in: McVey & Lee, 2008). Previously, it was thought that this alternative end-joining pathway was a “back-up” to the classical NHEJ pathway described above (Lieber *et al.*, 2004; Wang *et al.*, 2003). It has now been shown that MMEJ can play a prominent role in V(D)J recombination (Corneo *et al.*, 2007) and in Ig class switch recombination (Yan *et al.*, 2007). The MMEJ pathway is independent of NHEJ core factors such as Ku70/80 (Boulton & Jackson, 1996; Liang *et al.*, 1996; Ma *et al.*, 2003) and independent of HR core factors such as Rad51 and RAD52 (Lee & Lee, 2007). All the proteins involved in the MMEJ pathway as not known but certain factors have been identified in *S. cerevisiae* such as the MRE11 complex (Ma *et al.*, 2003), Nej1, and Srs2 (Lee & Lee, 2007) which are known to also have roles in either NHEJ pathway or single strand annealing pathway (Section 1.5.3).

It has also been reported that the Ku70 and Ku80 are not involved in VSG switching in *T. brucei*, as there was no detectable influence of the Ku mutants on the frequency of VSG switching (Conway *et al.*, 2002a). Moreover, it was found that end-joining using *T. brucei* cell extracts *in vitro* occurred using the pathway of MMEJ while there was no detectable NHEJ (Burton *et al.*, 2007). Interestingly, attempts in *T. brucei* as well as other *Kinetoplastida* have been unable to identify some components of the NHEJ machinery: DNA Ligase IV and XRCC4 (Burton *et al.*, 2007). Potentially, therefore, NHEJ is absent in these organisms and as end-joining DSB repair proceeds by MMEJ. This suggestion may be supported by the observation that the MMEJ pathway was found to be the dominant end-joining pathway for induced DSBs in *T. brucei* (Glover *et al.*, 2008). MMEJ has also been implicated in mammalian chromosome translocation, which may lead to cancer (Bentley *et al.*, 2004), and telomere dynamics in *Arabidopsis* (Heacock *et al.*, 2004).

## 1.5 Homologous Recombination (HR)

Homologous recombination (HR) allows repair of DSBs by copying a template that has the same DNA sequence, thereby allowing accurate duplication of the sequence at the DSB and preservation of the genomes integrity by ensuring chromosome stability (reviewed in: Li & Heyer, 2008; Otero & Hsieh, 1995; San Filippo *et al.*, 2008; Sung *et al.*, 2000; Symington, 2002; Wyman *et al.*, 2004). HR repairs DSB that have resulted from stalled or collapsed replication forks and exposure to DNA damaging agents (McGlynn & Lloyd, 2002; Thompson & Schild, 2001). HR also has a role in exchange of genetic information between maternal and paternal alleles during the generation of haploid gametes (meiosis), generating genetic diversity (Richardson *et al.*, 2004). HR has been shown to have a role in switching of mating types in yeast and in antigenic variation of pathogens, such as *T. brucei*.

### 1.5.1 Mechanism of Homologous Recombination: Gene conversion

The pathway of HR can be broken down into three stages, the presynaptic, the synaptic and the post-synaptic (Symington, 2002). In the presynaptic stage, the DSB is recognised and the DSB ends are processed to reveal 3' single stranded overhangs (ssDNA) which then allows the recombinase Rad51/RadA/RecA (in eukaryotes/archaea/eubacteria respectively) to bind to the DNA, with the aid of co-factors (Figure 1-11 A-D) (Benson *et al.*, 1994; Ogawa *et al.*, 1993; Sandler *et al.*, 1996). The synaptic or strand exchange stage occurs once a recombinase-ssDNA nucleoprotein filament has formed on the 3' DSB tails; the broken strand invades a homologous duplex DNA molecule where a D-loop is formed at the point of base-pairing (Figure 1-11 E-F) (Symington, 2002). The post-synaptic stage is where resolution of the strand exchange intermediates occurs, and can involve branch migration and resolution of crossed DNA strands (Figure 1-11 G-H) (Symington, 2002). These three stages appear to be broadly conserved in HR pathways of all kingdoms of life, where the orthologous recombinases (Rad51/RadA/RecA) appear to act in a strikingly similar manner (Benson *et al.*, 1994; Ogawa *et al.*, 1993; Sandler *et al.*, 1996). However, HR can occur by multiple pathways: gene conversion with or without a crossover event, break-induced replication and single strand annealing (reviewed in: Paques & Haber, 1999).

In eukaryotes the genes involved belong to an epistasis group initially defined in *S. cerevisiae* which include Rad50, Rad51, Rad52, Rad54, Rad55, Rad57, Rad59, Mre11, and

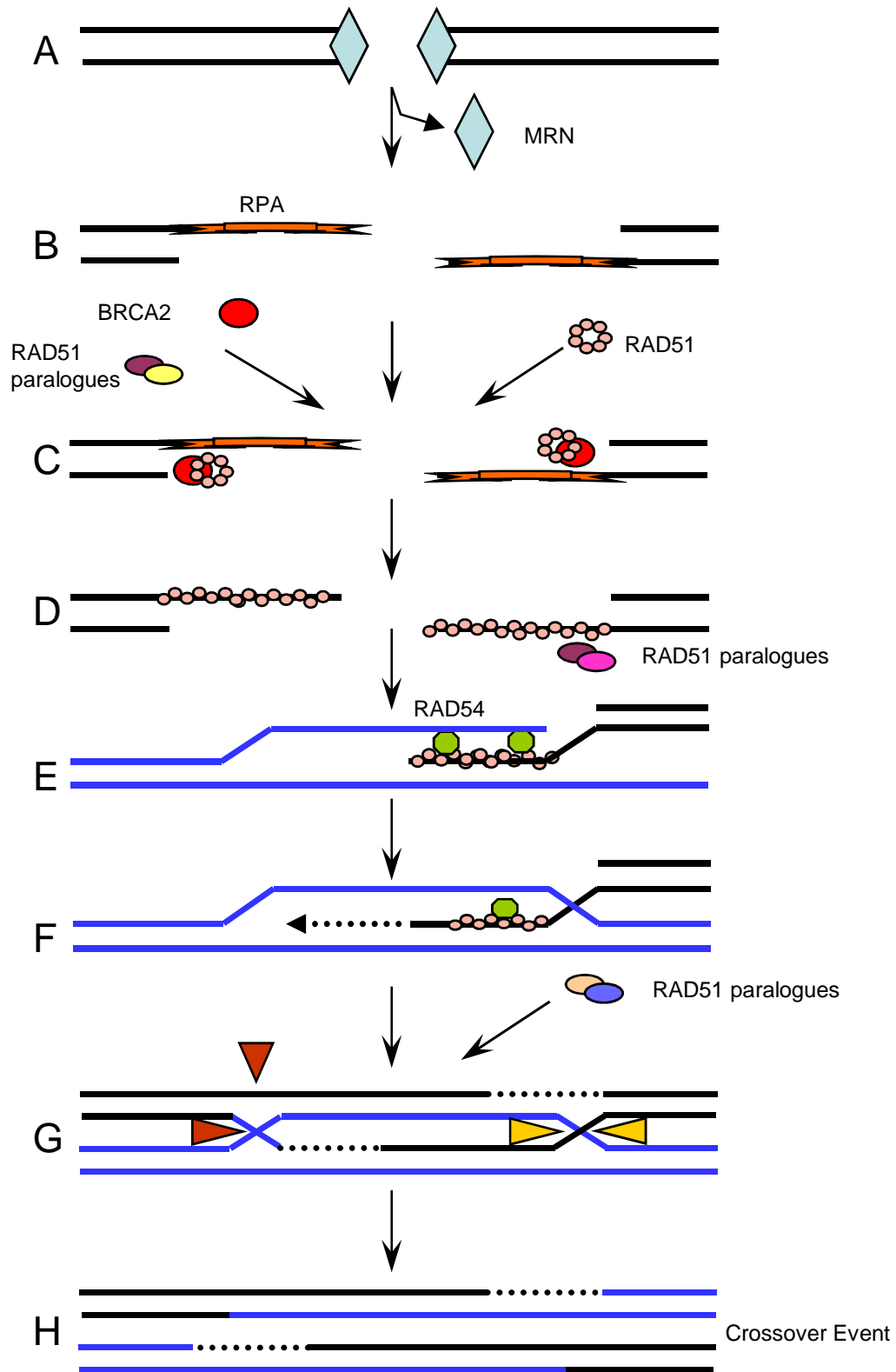


Xrs2 (Symington, 2002). Many of these factors are conserved proteins in most eukaryotes, and some in all the kingdoms of life. Recombination events must begin with a break in either one or both strands of the DNA helix, see Figure 1-11A. Once DSBs have occurred, the break ends are processed to reveal 3' ssDNA, see Figure 1-11A. Genetic studies have revealed that in *S. cerevisiae*, the protein complex comprised of Mre11, Rad50 and Xrs2 (the MRX complex; also called the MRN complex in mammals) has helicase and nuclease activities similar to the RecBCD complex, which aids this action in prokaryotic cells (Mimitou & Symington, 2008; Symington, 2002). The MRX complex also appears to function in the yeast NHEJ pathway, and similar to its role in HR, may be involved in processing of the break ends by nuclease activity although the enzymatic process of the MRX complex may differ for the two pathways (Dudasova *et al.*, 2004; Mimitou & Symington, 2008).

Replication protein A (RPA) binds with a high affinity to the free ssDNA, stimulates binding of Rad51 at the DSB and removes secondary DNA structures which could prevent effective base-pairing during strand exchange (Figure 1-11B and Figure 1-11C) (Baumann & West, 1998). The Rad51 protein wraps around the 3' ssDNA and forms a helical nucleoprotein filament in which one Rad51 molecule binds to every three nucleotide bases, in the presence of ATP; this is discussed in more detail in Section 1.6 (Figure 1-11D) (Benson *et al.*, 1994; Conway *et al.*, 2004). The nucleoprotein filament is needed for strand invasion of an intact DNA duplex or binding to another single strand of DNA and facilitates base pairing in this process, see Figure 1-11E (Chen *et al.*, 2008). Rad51 binding and nucleoprotein filament formation occurs with the help of a number of co-factors including a variable number of Rad51 paralogues (discussed in more detail, see Section 1.7), BRCA2, and Rad52, see Figure 1-11C-D. The detailed roles of these factors are still being elucidated (San Filippo *et al.*, 2008).

It has been shown that *S. cerevisiae* Rad52 helps to displace the RPA from the ssDNA, facilitating Rad51 binding (New *et al.*, 1998; Shinohara & Ogawa, 1998). *S. cerevisiae* Rad52 appears to have a role in all pathways of HR (Davis & Symington, 2004; Ivanov *et al.*, 1996) and its mutation has a strong impact on radiation sensitivity and Rad51 foci formation (Gasior *et al.*, 1998; Hays *et al.*, 1995). Most likely, this is because Rad52 can also promote some DNA strand exchange and can therefore allow Rad51-independent recombination (Mortensen *et al.*, 1996; Pohl & Nickoloff, 2008; Sugiyama *et al.*, 1998). In contrast, mutation of mammalian Rad52 has little effect on radiation sensitivity, Rad51 foci formation and has little effect on recombination rate (Rijkers *et al.*, 1998; van Veelen *et al.*, 2005; Yamaguchi-Iwai *et al.*, 1998). Nevertheless, mammalian Rad52 does appear to

form Rad52 foci when induced by DNA damage (Liu & Maizels, 2000). Interestingly, Rad52 is not conserved in all organisms and has not been detected in the trypanosomatids (Berriman *et al.*, 2005; El-Sayed *et al.*, 2005). In these organisms, the breast cancer type 2 susceptibility protein (BRCA2) may function in a similar manner to Rad52. Mutation in the BRCA2 gene of *T. brucei*, nematode and vertebrate cells results in sensitivity to DNA-damaging agents and defects in HR (Hartley & McCulloch, 2008; Martin *et al.*, 2005; Moynahan *et al.*, 2001). BRCA2 interacts directly with Rad51 (Pellegrini *et al.*, 2002; Yang *et al.*, 2002). It has also been shown to bind to ssDNA and dsDNA, displacing RPA and aiding Rad51 binding (Martin *et al.*, 2005; Yang *et al.*, 2002). These studies may provide evidence that BRCA2 functions in a similar manner to Rad52 in Rad52-deficient cells. However, it is unclear why mammalian cells have both Rad52 and BRCA2, and why there appears to be many other mediators needed for this step in HR. For example, two Rad51 paralogues have also been shown to displace RPA which aids the loading of Rad51 (Sigurdsson *et al.*, 2001) (see Section 1.7 below).



**Figure 1-11: Model for the homologous recombination DNA repair pathway.**

The diagram depicts one example of HR pathway, with the formation of two Holliday junctions leading to gene conversion with a crossover event. **A:** The DNA break ends are processed to reveal 3' ssDNA. **B:** Replication Protein A (RPA) binds to the ssDNA. **C:** Mediators help Rad51 binding. **D:** Rad51 forms a nucleoprotein filament. **E:** With the aid of Rad54, Rad51 catalyzes homology-based strand exchange of the single stranded end (Black lines) into the intact duplex DNA (Blue lines). The DSB can then be repaired by different pathways. Holliday junction (cross lines) formation and a crossover event are depicted here. **F:** The donor strand is displaced and copied. **G:** Resolving enzymes cleave along different axes for a crossover event to occur (Red and Yellow arrows) (Bhattacharyya *et al.*, 2004; Symington, 2002).

It has been shown that Rad54, a member of the Swi2/Snf2 protein family, associates with the Rad51 nucleoprotein and homologous double stranded DNA, and probably acts in facilitating the strand exchange by promoting homologous DNA pairing (Petukhova et al., 1998) (see Figure 1-11 E-F). It has also been suggested that as members of the Swi2/Snf2 family, Rad54, may function in chromatin remodelling when strand exchange occurs (Alexiadis & Kadonaga, 2002; Zhang et al., 2007). Rad54 is present in *T. brucei*, but has not been examined functionally. Once the nucleoprotein has invaded the duplex DNA strands and has found homologous regions, a so-called D-loop is formed (Symington, 2002), see Figure 1-11 F. Completion of HR pathway is now dependent on DNA polymerase extension from the 3' end of the invading strand; DNA synthesis occurs using the 3' end as a primer for new DNA synthesis and the donor strand as a template (Maloisel et al., 2004; Wang et al., 2004). One HR pathway involves the model proposed by Szostak with the formation of two Holliday junctions, four-stranded branched structures after the strand of the donor duplex is displaced and copied (Paques & Haber, 1999; Szostak et al., 1983). The resolution of these Holliday junctions can either lead to a cross over event or a non crossover event, depending on the way the strands are cut by a resolvase (Ip et al., 2008). If the non-crossover strand and cross over strand of both Holliday junction are cleaved in the same way, along the same axis, a crossover event will not take place, for example the yellow arrows on Figure 1-11 G. If the non-crossover strand and crossover strand of one junction are cleaved and the non-cross over and crossover strands of the other junction are cut along different axis, a crossover event will occur, for example the red arrows on Figure 1-11 G (Paques & Haber, 1999).

It would be predicted that resolution of Holliday junctions formed during HR would lead to an equal generation of crossover and non-crossover events, if the resolution was a random process. This is not the case, as *S. cerevisiae* mitotic recombination rarely results in a crossover event (Kupiec & Petes, 1988). As a result, a second pathway called the synthesis-dependent strand annealing (SDSA) model was proposed (Nassif *et al.*, 1994). SDSA occurs where one strand invades the DNA duplex, forming a D loop, and initiates DNA synthesis while the other strand remains unengaged (Paques & Haber, 1999). The newly synthesised strand from the D loop is then displaced from the template and anneals to the other 3' ssDNA tail, allowing DNA synthesis and ligation of the gaps (Haber *et al.*, 2004). It is proposed that SDSA occurs normally in non cross-over events but it can be modified to provide a crossover event as well (Ferguson & Holloman, 1996). Given the lack of gene crossover, it has been proposed that SDSA pathway may be a predominant mechanism for VSG switching (Barry, 1997; Borst *et al.*, 1996).

## 1.5.2 Break induced replication

Some significant variations on the relatively conserved HR pathway described above have been proposed (Section 1.5.1). DSB repair can also occur through a process called break induced replication (BIR) (Kraus *et al.*, 2001: see Figure 1-9C). This pathway can be Rad51-dependent or independent but appears to be Rad52-dependent (Ira & Haber, 2002; Signon *et al.*, 2001). In BIR, the DSB ends are also processed to reveal 3' ssDNA, but then only one end of the DNA break invades the duplex DNA and a true replication fork is formed which can copy a large portion, if not all, of the chromosome (McEachern & Haber, 2006). The second end of the DSB does not invade the duplex DNA, which can result in the loss of genetic information in that region (McEachern & Haber, 2006; Symington, 2002). It is uncertain the contribution that BIR has on the repair of DSB, although it has been proposed that BIR has a significant contribution in DSB repair in recombinase mutants (Malkova *et al.*, 2001) and potentially in telomere maintenance (McEachern & Haber, 2006). It has also been proposed that the BIR pathway plays a role in VSG switching for telomeric genes (Barry & McCulloch, 2001; Dreesen *et al.*, 2007).

## 1.5.3 Single strand annealing

Another non-conservative variant of HR is the mechanism of single strand annealing (SSA) (Lin *et al.*, 1990: see Figure 1-9B; Paques & Haber, 1999). This pathway can repair a DSB that is flanked by repetitive or homologous sequences. The DSB is processed by a 5'-3' exonuclease to reveal 3' ssDNA ends, which are complementary to one another (White & Haber, 1990). As both ends are homologous, this allows base pair annealing. Non-homologous DNA sequence is excised with loss of the genetic information (Fishman-Lobell *et al.*, 1992). Once the base pairs anneal, DNA synthesis follows and ligation repairs the gaps in the DNA strand. SSA appears to be dependent on Rad52 (Fishman-Lobell *et al.*, 1992; Sugawara & Haber, 1992) and can occur in *S. cerevisiae* *rad51*, *rad54*, *rad55*, and *rad57* mutants (Ivanov *et al.*, 1996).

## 1.5.4 Other roles of homologous recombination

### 1.5.4.1 Recombination and antigenic variation in *T. brucei*

In *T. brucei* cells, little is known about the specific pathways that regulate and act in HR in VSG switching (see Section 1.2.2). In bloodstream stage trypanosomes, VSG switching occurs by two mechanisms. >90% of switches appear to occur by DNA recombination, which delivers a previously silent VSG into a bloodstream expression site (BES).

Alternatively, the coat change can occur via transcriptional, or *in situ*, switching between different BESs (see Figure 1-5: Barry, 1997; Pays *et al.*, 2004). HR events in VSG switching primarily involves a non-reciprocal gene conversion, where the archive gene sequence is copied into the active BES (McCulloch, 2004). Following nuclease-mediated removal of the active VSG gene, DNA repair processes replace the lost gene by copying a silent VSG.

It is believed that the HR mechanisms, described above and in Figure 1-9, play a role in VSG switching, although there is a patchwork of different factors involved. It has been shown that mutants of *MRE11* in *T. brucei* exhibit chromosomal instability and impairment of HR. However *MRE11* does not contribute to recombination during antigenic variation in trypanosomes (Robinson *et al.*, 2002), perhaps implying that another nuclease may be involved in the VSG switching mechanisms, or one that plays a more substantial role than *MRE11*. Mutation of *RAD51*, the central enzyme of HR, impairs, but does not completely disrupt VSG switching, demonstrating that there are *RAD51*-dependent and *RAD51*-independent mechanisms involved (Conway *et al.*, 2002b; McCulloch & Barry, 1999). Given that the mediation of homology recognition and initiation of strand exchange requires the action of additional proteins (Baumann *et al.*, 1996), identification of the *RAD51* paralogues may explain or perhaps control some of the *RAD51*-independent pathways of VSG switching. Studies of two of these *RAD51*-like genes (*RAD51* paralogues) in VSG switching have shown that mutation of *RAD51-3* causes impairment of switching rates, while another, *RAD51-5* does not seem to be involved (Proudfoot & McCulloch, 2005a), suggesting that *RAD51* paralogues may have distinct roles in the process. Studies in the bloodstream stage demonstrated that mutation of a putative meiosis-specific recombinase, *DMC1*, does not affect HR, repair nor VSG switching (Proudfoot & McCulloch, 2005b). Finally, it has also been shown that *T. brucei* *BRCA2*, which is thought to act as a co-factor in HR, aiding *RAD51* loading onto ssDNA, has a role in VSG switching, as *brca2* *-/-* mutant cells have a decrease in VSG switching frequency of a very similar magnitude to the *rad51* *-/-* mutant cells (Hartley & McCulloch, 2008; McCulloch & Barry, 1999).

#### **1.5.4.2 Meiosis**

Meiosis is a process that results in gametogenesis and allows genetic variability of progeny. It is a specialised cell division that produces four haploid cells from one replicating parental diploid cell (Mange & Mange, 1999). The haploid cells have half the number of chromosomes compared to the parental cell. Chromosomal duplication in the

parental cell results in the formation of two identical chromatids, and is followed by chromosomal segregation and cell division (Mange & Mange, 1999). These two daughter cells divide again with the separation of each pair of chromatids into the progeny haploid cells. Fertilisation occurs by fusion of two haploid cells which restores the diploidy (Mange & Mange, 1999).

One important factor in meiosis is the arrangement of the recently replicated chromosomes into homologous pairs. This alignment of homologous chromosomes occurs in the stage of meiosis I called prophase I, and is characterised by synaptonemal complex (SC) assembly, genetic recombination and formation of chiasmata (Richardson *et al.*, 2004; Roeder, 1997). In prophase I, there are five main stages: leptotene, zygotene, pachytene, diplotene and diakinesis (Roeder, 1997). In leptotene there is the development of a synapse where the pair of chromosomal homologues become intimately associated with one another by a proteinaceous structure, the synaptonemal complex (Mange & Mange, 1999; Roeder, 1997). The SC is a ribbon-like structure that forms on the chromosomes and strengthens their synapsis (Mange & Mange, 1999). For a cross-over event to occur a DSB must form and strand exchange between the homologues occurs in the pachytene stage (Mange & Mange, 1999; Roeder, 1997). The cross-over event results in the exchange of genes between chromosomal homologues. Whether or not genetic exchange has occurred in the meiotic process is not physically evident until the diplotene stage as the chiasmata form (Mange & Mange, 1999). Chiasmata are visible under a light microscope for many cells and are where the two homologous, non-sister chromatids are connected. This is the location of the crossover event (Mange & Mange, 1999). The crossover events that occur during the process of meiosis enable the random shuffling of genetic information creating new combinations of genes. This allows an increase in the genetic diversity of the progeny compared to the parental cells, and strengthens the association between homologous chromosomes during meiosis.

The double strand break repair pathway similar to the one described in Section 1.4 allows meiotic recombination to occur (Richardson *et al.*, 2004). However, it is driven by some meiosis specific factors which have been described in different species. It has been shown that the topoisomerase-like Spo11 transesterase initiates the meiotic recombination by inducing DSBs in *S. cerevisiae* (Keeney *et al.*, 1997; Lichten, 2001). The DSB occurs in the leptotene stage at locations on the chromatid that are determined by local chromatin and chromosomal structure (Roeder, 1997). At the DSB, the DNA is processed to reveal 3' ssDNA ends probably by the complex of exonucleolytic proteins, Rad50 and Mre11 (Neale *et al.*, 2005). On the ssDNA, a nucleoprotein filament forms and strand exchange occurs

which involve Rad51 and, additionally, a meiosis specific homolog of Rad51 called Dmc1 (Bishop *et al.*, 1992; Ogawa *et al.*, 1993). It has also been shown that Rad52 and Rhd54/Tid1 are required for meiotic recombination (Lao *et al.*, 2008; Shinohara *et al.*, 2000). The segregation and assortment of alleles during genetic exchange between two *T. brucei* different strains occurs when co-transmitted through the Tsetse fly and is thought to involve meiosis (MacLeod *et al.*, 2005; Turner *et al.*, 1990). Most of the components required for meiotic recombination have been retained in the *T. brucei* genome (Klingbeil *et al.*, 2007), and one of which, DMC1, has been studied (Proudfoot & McCulloch, 2005b). DMC1 mutation in bloodstream stage cells causes no observable DNA repair or recombination defect, and perhaps this result is consistent with a function during a life cycle stage in the Tsetse fly in which meiosis occurs: the epimastigote stage (Gibson *et al.*, 2008; Gibson, 1995; MacLeod *et al.*, 2007; Proudfoot & McCulloch, 2005b: see Section 1.1.2).

#### **1.5.4.3 The switching of *S. cerevisiae* mating type**

In the budding yeast, *S. cerevisiae*, there are two sexes or mating types, **a** and  **$\alpha$** . This defines whether or not individual cells can sexually reproduce with the same or opposite mating type. The mating type is determined by the active gene present in the mating type locus (MAT). On the same chromosome as the MAT locus, there are two alleles that contain the information of the mating type, called HML and HMR (Klar, 1987). The HML is the locus that usually contains unexpressed mating type  **$\alpha$**  information, and the HMR is the locus that usually contains unexpressed mating type **a** information (see review: Haber, 2006; Symington, 2002). After meiosis and sporulation, each spore divides producing a mother and a daughter cell without a change in mating type. During the second mitotic division, it is only the mother cell which can switch mating type before the next round of cell division (Haber, 1992). Switching of the mating type occurs by gene conversion of the active gene with the other silent gene located in the donor locus.

The product of the *HO* gene, an endonuclease, was found to control the frequency of mating type switching of the yeast cell. There are two *HO* alleles; the dominant allele causes the homothallic phenotype which has a high mating type switching frequency (can be one switch per cell division) (Weiffenbach & Haber, 1981). The recessive *ho* is responsible for the heterothallic phenotype which has a much lower mating type switching frequency (Weiffenbach & Haber, 1981). The switching occurs by gene conversion in a similar pathway of DSB repair, described in Section 1.5 (Strathern *et al.*, 1982). It was



shown that that MAT switching was lethal in *rad52* mutants, suggesting the process involved HR (Malone & Esposito, 1980).

The existence of the enzymes Spo11 and HO, which generate a DSB that induces meiotic recombination and MAT switching respectively, suggest a mechanism by which HR can be targeted to drive a particular reaction (Haber, 1992). However, no evidence has been provided of such enzymes being involved in inducing breaks for VSG switching in trypanosomes, and nothing is known about the initiation events in this reaction. It may be the case that DSB arise randomly in specific hot spots in BES of a few percent of all parasites. It is equally possible a distinct process for initiation occurs, such as that used for MAT switching in yeast.

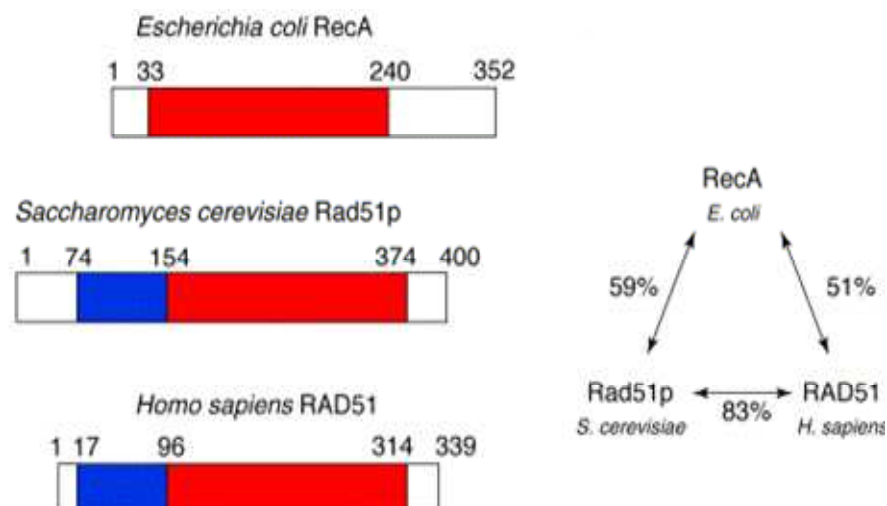
## **1.6 The role of the HR recombinases**

As discussed previously (Section 1.5.1), the genes involved in HR (HR) and DSB repair were identified primarily as radiation (rad) sensitive mutants in *E. coli*, *S. cerevisiae* and in rodent cells (Collins, 1993; Cox, 2000; Symington, 2002). In *S. cerevisiae*, the genes involved in DSB repair belong to the Rad52 epistasis group which include Rad51 and the Rad51 paralogues, and have some role in DNA strand exchange. Here, Rad51 and its homologues in other organisms are referred to as the recombinases.

### **1.6.1 Phylogeny of the recombinases and their role in strand exchange**

RecA is highly conserved among the eubacteria (Clark & Margulies, 1965). RecA is a multifunctional protein, with involvement in the induction of the SOS response to DNA damage in prokaryotes, through regulation of the LexA repressor protein (Little & Mount, 1982). RecA-mediated cleavage of LexA results in expression of many genes, most of which are involved in DNA repair pathways. Analysis of RecA also shows that it promotes the central steps of HR and can be defined as a DNA-stimulated ATPase (Craig & Roberts, 1981). RecA binds to ssDNA or partially single stranded DNA to form a nucleoprotein filament which can promote strand exchange. It has been shown that RecA binds with a higher affinity to ssDNA than dsDNA. The strand exchange-proficient RecA nucleoprotein filament forms in a 5'-3' direction relative to the single strand (Register & Griffith, 1985), and with ATP hydrolysis, the exchange occurs in an unidirectional manner (Jain *et al.*, 1994). This nucleoprotein filament is helical and consists of approximately six monomers of RecA per helical turn, with 2-3 bound bases per RecA monomer (Di Capua *et al.*, 1982).

Studies revealed that the prokaryotic RecA and its orthologs, RadA and Rad51 in archaea and eukaryotic respectively (Sandler *et al.*, 1996; Shinohara *et al.*, 1993), are conserved in all sequenced organisms, implying that they are essential for HR in living cells (Kawabata *et al.*, 2005). Analysis comparing *E. coli* RecA and its Rad51 ortholog in yeast *S. cerevisiae*, revealed that over 30 % amino acid sequence was identical when the core domains of each protein were compared (Baumann & West, 1998; the red domains of Figure 1-12). Sequence analysis also showed high homology of the archaean RadA, with RadA having 40 % identity to yeast Rad51 and around 20 % to RecA (Sandler *et al.*, 1996). The human Rad51 protein contains a core region that has 67 % identity to *S. cerevisiae* and 30 % identity with *E. coli* RecA (Baumann *et al.*, 1996; Shinohara *et al.*, 1993). These core regions contain two highly conserved consensus motifs called the Walker A and Walker B motifs, or boxes (Lin *et al.*, 2006). These motifs are conserved in ATPases and confer ATP binding and hydrolysis activities to the RecA/Rad51 proteins (Walker *et al.*, 1982). The nucleoprotein filament formed by *S. cerevisiae* Rad51 was high analogous structurally and stoichiometrically to that formed by RecA, with the Rad51 filament binding 2-3 nucleotides per monomer (Benson *et al.*, 1994; Conway *et al.*, 2004; Ogawa *et al.*, 1993). Rad51, from yeast and mammals, was also shown to have a similar function to RecA, having a role in HR and DSB repair.



**Figure 1-12: Sequence comparisons of prokaryotic RecA with eukaryotic Rad51 from yeast and humans.**

Schematic representation of the domains of the three proteins and their relative amino acid sizes; *E. coli* RecA, *S. cerevisiae* Rad51 and *H. sapiens* Rad51. The conserved core domains are in red and the N-terminal sequences conserved in *S. cerevisiae* and *H. sapiens* RAD51 are in blue. The averages of the identity and similarity percentages of the amino acids residues within the core domains are shown. Reprinted with permission from Elsevier: Baumann *et al.* (1998), Trends in Biochemical Sciences, 23, pp 248.

Despite these structural similarities and high identity between the core domains of the proteins, there are some other differences between RecA and Rad51. The proteins differ

significantly outside the core domain (Figure 1-12); The RecA and Rad51 proteins have variable N and C terminal domains (the white domains of Figure 1-12). The N-terminal region of Rad51 and RadA (not shown) have a Helix-hairpin-Helix (H-h-H) domain, which appears to be absent from RecA and may have a role in DNA binding (Aihara *et al.*, 1999). The prokaryotic RecA and *Arabidopsis thaliana* RecA1 have a C-terminal domain absent in orthologous Rad51 or RadA proteins. This C-terminal domain of RecA has been shown to bind to dsDNA, although the function is unknown (Aihara *et al.*, 1997). It may have a role in strand exchange due to its dsDNA binding properties and maybe even be important for the packing of adjacent polymers in the RecA-DNA filament (Aihara *et al.*, 1997; Story *et al.*, 1992).

These structural differences between Rad51 and RecA could account for some observed functional differences. Firstly, the rate of ATP hydrolysis of Rad51 is much lower than that of RecA. Secondly, while RecA and Rad51 have been shown to form a filament on nicked duplex DNA and ssDNA in the same manner, it has been shown that human Rad51 can bind highly supercoiled DNA *in vitro*, while RecA was limited by the torsional stress caused by extensive unwinding (Benson *et al.*, 1994). Thirdly, the DNA binding preferences of RecA and Rad51 appear to be different. RecA prefers to bind to single stranded or partially single stranded DNA, while the binding affinity of Rad51 proteins to ssDNA and dsDNA are similar. Fourthly, in strand exchange, the Rad51 filament forms on the complementary strand in the 3'-5' direction which is the opposite direction to that of RecA filament (Sung & Robberson, 1995). Finally, RecA has been shown to form as a large nucleoprotein filament across the DNA, while this is not the case for Rad51 (van der Heijden *et al.*, 2007). The Rad51 nucleoprotein formation consists of many small filament fragments that are only a few tens of monomers long (Galletto *et al.*, 2006). This high flexibility and dynamic nature of this arrangement is likely to facilitate strand exchange of Rad51 (Galletto *et al.*, 2006).

Most eukaryotes possess a second recombinase, called Dmc1 (Disrupted meiotic cDNA protein 1), which potentially functions exclusively in meiosis and is not found in prokaryotes (Bishop *et al.*, 1992). *S. cerevisiae* Dmc1 was shown to be highly homologous to Rad51 since 45 % of the amino acid residues are identical between *S. cerevisiae* Dmc1 and Rad51. *S. cerevisiae* Dmc1 was also shown to have homology with *E. coli* RecA: with 26 % identity and 41 % similarity of the amino acid residues in the core regions. Dmc1 is required for normal levels of meiotic recombination and progression (Bishop *et al.*, 1992). *H. sapiens* Dmc1 was shown to bind to ssDNA, and have a weak ATPase activity compared to RecA. This study also concluded that *H. sapiens* Dmc1 has a role in strand

exchange (Li *et al.*, 1997). Interestingly, *Caenorhabditis elegans* does not have Dmc1 however, and the Ce-Rad51 (Ce-rdh-1) protein has been shown to have both Rad51 and Dmc1 properties, with Ce-Rad51 being 59 % identical to *Mus musculus* Rad51 and 51 % identical to *M. musculus* Dmc1. The Ce-Rad51 was also shown to be involved in DSB repair in mitotic and meiotic nuclei, and to act in meiotic progression (Takanami *et al.*, 1998).

## 1.6.2 Phenotypes of recombinase mutants

Studies have revealed that mutation of *E. coli* *RecA* results in an increased sensitivity to radiation, and a decrease in HR rates (Clark & Margulies, 1965). As with *E. coli*, *S. cerevisiae* *rad51* mutants are viable and are observed to be defective in recombination as well as having an increased sensitivity to a DNA damaging agents, including MMS (Chanet *et al.*, 1996). A *RadA* mutant has been generated in the archaeal extreme halophile, *Haloferax volcanii* (Woods & Dyll-Smith, 1997). This mutant has a slower growth rate than wildtype cells, had increased sensitivity to UV damage and DNA recombination was not detectable (Woods & Dyll-Smith, 1997). The T4 bacteriophage recombinase, UvsX, was first defined by mutation in “gene 49” that caused the phage to become sensitive to UV radiation, have a decreased level of genetic recombination, a decreased burst size, and an inability to replicate mature DNA (Miyazaki *et al.*, 1983; Yonesaki *et al.*, 1985). Mutated *rad51* in *T. brucei* cells are viable (McCulloch & Barry, 1999), though such cells display phenotypes such as susceptibility to DNA damage with DNA damaging agents, MMS and phleomycin, and impaired HR, see Chapter 3 and Chapter 4 of this thesis.

The above information leads to the realisation that RecA/Rad51/RadA proteins are crucial in HR but are not essential. In contrast, generation of conditional *rad51* mutants in chicken DT40 cells showed chromosomal aberrations before cells underwent cell cycle arrest and death (Sonoda *et al.*, 1998). Similarly, a fatal mutation of the *rad51* gene in embryonic cells indicated that mammalian Rad51 is essential for proliferation and survival during embryogenesis (Lim & Hasty, 1996; Tsuzuki *et al.*, 1996). Knockouts of the meiotic-specific protein, *dmc1*, in embryonic mice are viable but are infertile.

## 1.7 The roles of the eukaryotic RAD51 paralogues

As discussed above, the recombinases are conserved throughout evolutionary but some differences between Rad51 and RecA have been described (see Section 1.6.1). These biochemical and structural differences may be explained, at least in part by the evolution of co-factors that aid the function of the recombinases. Such co-factors are found in different

numbers in different organism, and appear only to be semi-conserved. Prominent amongst this group of putative recombinase co-factors are the so-called Rad51 paralogues, Rad51-like genes that are found in variable numbers of eukaryotes (see Chapter 7, Section 7.2.2).

### 1.7.1 Identified Rad51 paralogues

The identification and analysis of Rad51-like factors, the so called Rad51 paralogues, has suggested distinct functions of this set of proteins during HR. To date, no bacterial RecA paralogues have been identified, suggesting they are specific to eukaryotes and archaea. Within their Kingdoms, the roles and properties of the Rad51 paralogues have been best described in vertebrates and in yeast. *H. sapiens* and mammals in general, have five RAD51 paralogues named Rad51B (also called Rad51L1/HsRec2), Rad51C (Rad51L2), Rad51D (Rad51L3), XRCC2 and XRCC3 (Schild *et al.*, 2000). Like mammalian cells, *Arabidopsis thaliana* has five Rad51 paralogues, which are named in the same way as the mammalian proteins, reflecting their likely orthology (Bleuyard *et al.*, 2005). In contrast, the budding yeast, *S. cerevisiae*, has only two Rad51 paralogues, called Rad55 and Rad57. While the fission yeast, *Schizosaccharomyces pombe*, has four more recently identified paralogues called Rhp55, Rhp57, Rlp1 and Rdl1 (Khasanov *et al.*, 2004; Martin *et al.*, 2006; Tsutsui *et al.*, 2001). Unusually, in some species only one Rad51 paralogue has been identified. For instance, the nematode *Caenorhabditis elegans* has RFS-1 (Boulton *et al.*, 2002; Ward *et al.*, 2007), the phytopathogenic fungus *Ustilgo maydis* has Rec2 (Holloman *et al.*, 2008) and the archaea have RadB (Guy *et al.*, 2006); in all cases these proteins have been shown to have a role in DNA repair. In *Drosophila melanogaster*, where the Rad51 homologue is called SpnA, four SpnA paralogues are found, SpnB, SpnD, CG2412, CG6318, although only two of these have been functionally characterised (Abdu *et al.*, 2003). Even within the Kinetoplastida, the numbers of Rad51 paralogues are not conserved. Four RAD51 paralogues have been identified in *T. brucei* named RAD51-3, RAD51-4, RAD51-5, RAD51-6 and as well as the meiotic-specific protein DMC1 (Proudfoot & McCulloch, 2005a), and each can be identified in *T. cruzi*. *Leishmania major* however has only three RAD51 paralogues and appears to lack the RAD51-5 (Proudfoot, 2005). The varied number of the Rad51 paralogues throughout the major phylogenetic groups and the loss of RAD51-5 from *Leishmania* will be considered further in the discussion (see Chapter 7, Section 7.2.2).

### 1.7.2 Phenotypes of *rad51* paralogue mutants in yeast

In *S. cerevisiae* null mutations of the two Rad51 paralogues, *rad55* and *rad57*, have been described and each mutant causes cold sensitivity for DNA repair and recombination

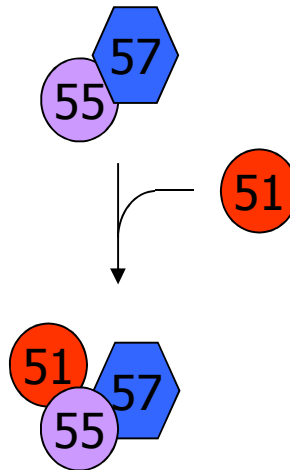
(Game & Mortimer, 1974; Hays *et al.*, 1995; Lovett & Mortimer, 1987). A mutation in the conserved lysine residue within the Walker A motif in Rad55 affects recombination, whereas an equivalent mutation in Rad57 has no effect (Johnson & Symington, 1995). Rad55 and Rad57 are both required for mating-type switching in the yeast, which is a specialised mitotic recombination event (Johnson & Symington, 1995; Sugawara *et al.*, 1995; see above Section 1.5.4.3). *rad55* and *rad57* mutants also result in a decrease in DNA damaged-induced Rad51 foci formation (Gasior *et al.*, 1998), consistent with a role in the Rad51-mediated strand exchange stage in HR.

As discussed above, *S. pombe* has Rhp55 and Rhp57 proteins that are orthologs of Rad55 and Rad57 in *S. cerevisiae*, respectively, as well as two additional Rad51 paralogues named Rlp1 and Rdl1 (Khasanov *et al.*, 2004; Martin *et al.*, 2006; Tsutsui *et al.*, 2001). However, *S. pombe rhp55* and *rhp57* mutants display phenotypes that are different to that of *rad55* or *rad57* budding yeast mutants. For example, a mutation of the conserved lysine residue of the Walker A motif in Rhp55 and Rhp57 affected both cell lines when treated with the DNA-damaging agent, MMS (Tsutsui *et al.*, 2001). Also, the *S. pombe rhp55* and *rhp57* mutants were more sensitive to MMS at lower temperatures compared to the *S. cerevisiae rad55* and *rad57* mutants (Johnson & Symington, 1995; Tsutsui *et al.*, 2001). Rhp55 and Rhp57 have roles in spore viability, and all four fission yeast Rad51 paralogues have roles in meiotic recombination (Grishchuk & Kohli, 2003).

### 1.7.3 Interactions and roles of yeast RAD51 paralogues

*S. cerevisiae rad55* and *rad57* mutants display cold sensitivity, which is a property associated with proteins that complex as multiple subunits or with other proteins, indicating that the proteins act as a complex. This was shown to be the case, with Rad55 and Rad57 found to interact in yeast-two hybrid analysis (Hays *et al.*, 1995; Johnson & Symington, 1995) and that Rad55 and Rad57 form a stable heterodimer based on co-immunoprecipitation data (Sung, 1997). It has been shown that Rad51-mediated strand exchange can be enhanced by the addition of Rad55 and Rad57, *in vitro* (Sung, 1997). Rad55 also interacts with Rad51 in yeast-two hybrid analysis, while there is no evidence that Rad57 interacts with Rad51 (Hays *et al.*, 1995; Johnson & Symington, 1995; see Figure 1-13 ). In fission yeast, Rhp55 and Rhp57 also interact with one another but, in contrast to budding yeast, Rhp57 interacts Rad51 (Rhp51) (Tsutsui *et al.*, 2001). It is noteworthy (see below) that mammalian Xrcc3 and Rad51C also interact with Rad51, since these have higher homology with Rhp57 compared Rhp55 (Grishchuk & Kohli, 2003). *S. cerevisiae* Rad55 interaction with Rad51 has been shown to result in stabilisation

of the Rad51 nucleoprotein filament on the single-stranded DNA (Miyazaki *et al.*, 2004). Rad55-Rad57 have also been shown to overcome the inhibition of Rad51 strand exchange caused by RPA (Sung, 1997), and overcome the DNA repair and HR defects of *rad55* and *rad57* mutants (Hays *et al.*, 1995; Johnson & Symington, 1995). As a result of these analyses, it appears that the Rad55 and Rad57 paralogues interact as co-factors for the formation and stabilisation of Rad51 nucleoprotein filaments in DSB repair.



**Figure 1-13: Predicted model of complex formations and interactions among *S. cerevisiae* Rad51 and the Rad51 paralogues.**

*S. cerevisiae* Rad51 paralogues, Rad55 (55) and Rad57 (57), are thought to act as a heterodimer. Rad55 interacts with the *S. cerevisiae* recombinase, Rad51 (51), helping polymerisation of Rad51 on ssDNA, resulting in strand exchange and DSB repair.

#### 1.7.4 Phenotypes of *rad51* paralogue mutants in vertebrate cells

Two mammalian Rad51 paralogue proteins, XRCC2 and XRCC3, were identified by functional complementation of the *irs1* and *irs1SF* mutant hamster cell lines. These cell lines had common phenotypes: they were sensitive to DNA damaging agents, extremely sensitive to DNA cross-linking agents and displayed gross chromosomal aberrations. These were so-called the x-ray repair cross complementing (XRCC) genes as complementation of the XRCC2 and XRCC3 in *irs1* and *irs1SF* respectively efficiently restored cell function (Liu *et al.*, 1998). In contrast, Rad51B (Albala *et al.*, 1997), Rad51C (Dosanjh *et al.*, 1998) and Rad51D (Pittman *et al.*, 1998) were identified by sequence database analysis based on similarity to Rad51.

Mutational analysis in mice has revealed that *rad51B*, *rad51C*, *rad51D* and *xrcc2* mutations are embryonic lethal where the fetuses were small and malformed (compared with a healthy fetus) and died in early-mid gestation (Deans *et al.*, 2000; Kuznetsov *et*

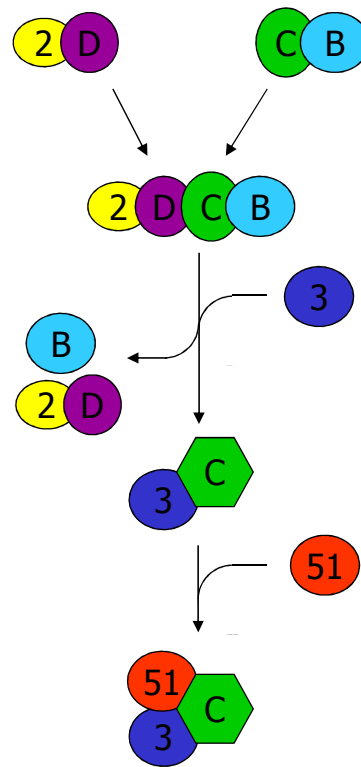
*al.*, 2007; Pittman & Schimenti, 2000; Shu *et al.*, 1999). Despite this lethality of embryonic cells, Rad51 paralogue somatic cell mutants, such as hamster cell lines (V79 and CHO) and chicken lymphocyte cell lines (DT40) are viable (Liu *et al.*, 1998; Takata *et al.*, 2000; Takata *et al.*, 2001). This allowed phenotypic analysis of the mutant cell lines and potential roles to be assigned to the Rad51 paralogues. Studies by the Takeda group showed that mutation of *rad51B*, *rad51C*, *rad51D*, *xrcc2* and *xrcc3* in chicken DT40 cell lines resulted in very similar phenotypes, with a decreased growth rate, an increased sensitivity to DNA damaging agents (both  $\gamma$ -rays (mild) and cross-linking agents (strong)), and a reduced ability to undergo HR relative to wildtype cells (Takata *et al.*, 2000; Takata *et al.*, 2001). These phenotypes were similar to the ones seen in the *irs1* and *irs1SF* mutants lacking in a functional XRCC2 and XRCC3 respectively (Liu *et al.*, 1998). It has also been observed that mutation of *xrcc2* and *xrcc3* affected HR rates compared to hamster wildtype cells (Johnson *et al.*, 1999; Pierce *et al.*, 1999). Rad51 paralogue mutants in chicken DT40 cells also showed a decrease in DNA damage-induced Rad51 foci formation (Takata *et al.*, 2000; Takata *et al.*, 2001). This was especially relevant, showing a role for the Rad51 paralogues in relocalisation of Rad51 into foci when DNA damage occurred or in stabilisation of the Rad51 foci (Haaf *et al.*, 1995).

It is interesting to note that the phenotypes seen in the Rad51 paralogue mutant chicken DT40 cell lines can be fully or partially overcome by over-expression of human Rad51 and with complementation of the corresponding human Rad51 paralogue gene (Takata *et al.*, 2000; Takata *et al.*, 2001). Also, two studies have shown that the early-mid embryonic lethality of Rad51B and Rad51D mutation in mice could be partially overcome by their generation in a *p53*  $-/-$  background. *p53* has a role in checkpoint cell cycle control and apoptosis, and loss of *p53* is sufficient for an increase in proliferation of the highly comprised Rad51B and Rad51D mutant fetuses (Shu *et al.*, 1999; Smiraldi *et al.*, 2005).

### **1.7.5 Interactions and roles of Rad51 paralogues in vertebrate cells**

Yeast-two hybrid and co-immunoprecipitation analyses have been carried out to examine the interaction between the mammalian Rad51 paralogue proteins. These experiments showed that Rad51 paralogues in human cell extracts form two discrete protein sub-complexes, one containing Rad51B-Rad51C-Rad51D-XRCC2 (BCDX2) and the other Rad51C-XRCC3 (CX3) (Masson *et al.*, 2001b; see Figure 1-14 ; Schild *et al.*, 2000). Later analysis supported these findings, confirming that Rad51C was indeed present in the two Rad51 paralogue complexes, BCDX2 and CX3 (Liu *et al.*, 2002b).





**Figure 1-14: Predicted model of complex formations and interactions among mammalian Rad51 and the Rad51 paralogues.**

Rad51, Rad51B, Rad51C, Rad51D, XRCC2 and XRCC3 are shown in the diagram as 51, B, C, D, 2 and 3 respectively. XRCC2-Rad51D and Rad51B-Rad51C dimers combine to form the Rad51B-Rad51C- Rad51D-XRCC2 complex. XRCC3 may promote the dissociation of Rad51C from this complex to produce the XRCC3-Rad51C dimer that binds to Rad51 (Liu *et al.*, 2002a).

Further analysis of the Rad51 paralogue complexes has attempted to determine their function. The first of these studies examined the sub-complex comprising of the two Rad51 paralogues, Rad51B-Rad51C (BC). As mentioned in Section 1.5.1 above, RPA is needed for Rad51-mediated strand exchange. RPA binds to the ssDNA and is thought to remove any secondary structures that may hinder the HR pathway. As RPA is bound to the ssDNA, it competes with Rad51 for binding sites to the ssDNA template. It was shown *in vitro* that this inhibitory effect of RPA could be overcome by the BC complex, with an increase of strand exchange products when BC complex was added to the reaction containing Rad51, RPA and ssDNA (Sigurdsson *et al.*, 2001). Another biochemical analysis focused on another Rad51 paralogue complex, Rad51C-XRCC3 (CX3). This suggested that CX3 had Holliday junction resolvase activity (Liu *et al.*, 2004). As described in Section 1.5.1, some recombination events require the formation of a Holliday junction intermediates, where recombining DNA strands cross. Resolution of these junctions in bacteria is carried out by RuvABC. It was found that *rad51C* and *xrcc3* mutant human cell extracts had reduced levels of resolvase activity, while addition of purified complexes with Rad51C such as the BC complex to the extract restored all activity (Liu *et al.*, 2004). Additionally, it has been shown that Rad51C stabilises XRCC3, as knocking down Rad51C levels in HeLa cells by

RNAi also greatly reduced XRCC3 levels (Lio *et al.*, 2004). The data above have been used to generate the interaction model in Figure 1-14, which suggest that the Rad51 paralogues form at least two distinct complexes, which potentially provide distinct functions within the mammalian cell.

### **1.7.6 Phenotypes of the *rad51-3* and *rad51-5* mutants in *T. brucei* cells**

Four RAD51 paralogues have been identified in *T. brucei*, RAD51-3, RAD51-4, RAD51-5, RAD51-6 and the meiotic-specific protein DMC1 (Proudfoot & McCulloch, 2005a). To date, mutations in only *RAD51-3* and *RAD51-5* have been generated. Both mutations result in HR impairment (but not complete abolishment) and phenotypic sensitivity to DNA damage (Proudfoot & McCulloch, 2005a). In addition, both mutants show a reduced capacity for Rad51 re-localisation into discrete foci following phleomycin DNA damage. Together, these preliminary data are consistent with the hypothesis that the two proteins act as RAD51 mediators (Proudfoot & McCulloch, 2005a) similar to the suggestion in yeast and mammals (this is discussed in Chapter 4). In contrast to this, studies in the bloodstream stage could find no evidence that *T. brucei* DMC1 acts in HR, DNA repair, or in antigenic variation (Proudfoot & McCulloch, 2005b).

### **1.7.7 Some examples of the phenotypes of the Rad51 paralogue mutants in other eukaryotic cells**

#### **1.7.7.1 The Rad51 paralogue in *Caenorhabditis elegans***

RFS-1 is the only identified *C. elegans* Rad51 paralogue and is closely related to human and mouse Rad51D (Boulton *et al.*, 2002; Martin *et al.*, 2005). RFS-1 was suggested not to have a role in meiosis as meiotic crossover events and meiotic Rad51 foci form in *rfs-1* mutants (Ward *et al.*, 2007). The phenotype of *rfs-1* mutants was found to be different to that seen in *C. elegans brca2* mutants, where there was a decrease in Rad51 foci formation (Martin *et al.*, 2005; Ward *et al.*, 2007). The *rfs-1* mutant was also moderately sensitive to irradiation-induced DSBs and severely sensitive to cross linking agents, such as cisplatin, compared with wildtype cells (Ward *et al.*, 2007). It was concluded that RFS-1 had a role in genome integrity primarily due to a role in replication as *rfs-1* mutants were severely sensitive to agents that are thought to impact on replication fork progression. Interestingly, successive generations of nutritionally starved *rfs-1* mutant cells results in progressive sterility which can be defined as Mortal germline mutants (Mrt). These Mrt mutants also

develop telomere erosion suggesting another role of the RFS-1 in genome stability (Yanowitz, 2008).

#### **1.7.7.2 The Rad51 paralogue in *Ustilago maydis***

The fungus *Ustilago maydis* also has only one Rad51 paralogue, Rec2. This is an unusually large protein relative to the other Rad51 paralogues, with an amino acid sequence of 782 which is twice the size of the *U. maydis* Rad51 protein. Comparisons between the *U. maydis* Rad51 and Rec2 amino acid sequences show that the Walker A and B boxes have 52 % and 21 % identity and 76 % and 65 % similarity, respectively (Ferguson *et al.*, 1997). Mutants of *rec2* have an increased sensitivity to DNA damaging agents (UV, MMS and  $\gamma$ -radiation) and recombination rates are also decreased. The phenotypes of *rec2* mutants was similar to that observed in the *U. maydis* Rad51 mutant cells (Ferguson *et al.*, 1997). Strikingly, analysis of the purified 110 kDa REC2 protein showed that this Rad51 paralogue is able to bind to ssDNA, has DNA-dependent ATPase activity and is able pair sequences that have homology (Kmiec *et al.*, 1994), activities not found for any other Rad51 paralogue.

### **1.8 Aims of the project**

The overall aim of this thesis was to examine further the roles of the *T. brucei* RAD51 paralogues in the DSB repair pathway of HR and, in particular, to assess whether or not this group of related proteins provide a single function or multiple functions. To this end, three approaches were adopted.

#### **1.8.1 Examining the role of RAD51-4 and RAD51-6**

To date, generation and analyses of *rad51*, *rad51-3*, *rad51-5*, and *dmc1* mutants have been conducted (McCulloch & Barry, 1999; Proudfoot & McCulloch, 2005a; Proudfoot & McCulloch, 2005b). It was therefore necessary to examine the functions of the two remaining RAD51 paralogues, RAD51-4 and RAD51-6, to build a complete picture of the possible roles of these genes in *T. brucei*. If *T. brucei* RAD51-4 and RAD51-6 have similar roles to RAD51-3 and RAD51-5 and potentially to Rad51 paralogues of other species, it would be predicted that the proteins will contribute to general HR and DNA damage repair. It is difficult to predict whether or not RAD51-4 and RAD51-6 will be involved in VSG switching, given the conflicting data on RAD51-3 and RAD51-5, see Chapter 4, Section 4.6. To this end, heterozygous and homozygous mutants of the *T. brucei* RAD51-4 and RAD51-6 genes were generated by gene disruption. This involved

replacing the open reading frame (ORF) of the gene with one of two antibiotic resistance cassettes, puromycin and/or blasticidin. Phenotypic analyses of the confirmed mutants were carried out: growth rates, sensitivity to DNA-damage agents, transformation efficiencies, ability to form DNA damage induced RAD51 foci and VSG switching were analysed (Chapter 3). These phenotypes were then compared to those observed in the *rad51-3* *-/-* and *rad51-5* *-/-* cells (Chapter 4).

## **1.8.2 Analysing the association between the RAD51-paralogues**

To determine if the RAD51 paralogues form discrete sub-cellular complexes similar to that of the Rad51 paralogues in mammals, or a single complex analogous to yeast Rad55/Rad57, interaction between the proteins was assessed. Two approaches were used (Chapter 5). The first approach involved yeast-two-hybrid assays. For this, the genes were cloned into bait and hunter plasmids, pHybLex/Zeo and pYesTrp, respectively (Invitrogen). Western analysis was used to show protein expression and analysis of any interaction(s) was performed using the Yeast  $\beta$ -galactosidase Assay kit (Pierce), X-gal filter lift assay, and histidine auxotrophy analysis. In the second approach, immunoprecipitation of the RAD51 paralogues was attempted to determine interactions between the proteins within *T. brucei* cells. To aid testing for protein-protein interactions, anti-sera was generated against the RAD51 paralogues and used for co-immunoprecipitation. Where such anti-sera was found to be inappropriate due non-specific binding, *T. brucei* strains were generated in which the RAD51 paralogues were expressed with an epitope tag, PTP (Schimanski *et al.*, 2005a; Schimanski *et al.*, 2005b). This tag allows immunoprecipitation of the proteins which can be followed by western analysis.

## **1.8.3 A luciferase assay for *T. brucei* homologous recombination**

The third experimental approach of this thesis was to design a novel assay to test for HR rates and pathways in *T. brucei*. To date, all general HR in *T. brucei* has been assayed by a relatively artificial route, evaluating the ability of the parasite to undergo HR by measuring the rate of integration of a construct containing an antibiotic resistance marker. This assures that the number of antibiotic resistant transformants will directly relate to the efficiency of HR. In fact, this assays, measures transformation rather than recombination, and may be prone to experimental fluctuation. The objective of this approach was to develop an alternative assay, which is more physiological and adaptable (Chapter 6).

## **Chapter 2. Materials and Methods**

## 2.1 Cell culturing

### 2.1.1 Trypanosome growth

*T. brucei* bloodstream stage 3174.2, a derivative of the strain Lister 427, was used and analysed in all the experiments described in this thesis. Lister 427 strain is primarily monomorphic, and has adapted to life in laboratory conditions and has the ability to infect rodents (Melville *et al.*, 2000). The derivative of Lister 427 strain, 3174.2, has two resistance markers, (neomycin and hygromycin resistance cassettes) in the bloodstream expression site that expresses MITat1.2a, VSG221. Therefore the 3174.2 strain enables the VSG switching frequencies and mechanisms to be analysed, and is discussed in more detail in Section 3.4.5 (McCulloch *et al.*, 1997; Rudenko *et al.*, 1996).

All *T. brucei* bloodstream stage strain 3174.2 cells were grown in HMI-9 medium (Hirumi & Hirumi, 1994). This media contains a powder supplied by Invitrogen containing hypoxanthine, L-cysteine, pyruvic acid, thymidine, bathocuproedisulfonic acid and  $\beta$ -mercaptoethanol. The medium was made by addition of tissue culture water (Sigma) to the sachet of powder (Invitrogen) and completed with 10 % (v/v) fetal bovine serum and 1 % penicillin/streptomycin (Invitrogen). The cells were passaged on a twice/thrice weekly basis and grown in a 37 °C, 5 % CO<sub>2</sub> incubator.

### 2.1.2 Transfecting of *T. brucei*

*T. brucei* bloodstream stage strain 3174.2 cells were grown in HMI-9 medium (see above). An aliquot of  $5 \times 10^7$  cells was centrifuged at 600 x g for 10 minutes at room temperature. The cells were then resuspended in 0.5 ml of prewarmed Zimmerman Medium with glucose [132 mM NaCl, 8 mM KCl, 8 mM Na<sub>2</sub>HPO<sub>4</sub>, 1.5 mM KH<sub>2</sub>PO<sub>4</sub>, 0.5 mM Mg acetate, 0.09 mM Ca acetate, 0.25 mM D-glucose (Sigma)]. Cells were electroporated in 0.2 cm cuvettes (Bio-Rad) with approximately 5  $\mu$ g of plasmid DNA at 1.5 kV and 25  $\mu$ F capacity. The construct DNA had been linearised using restriction enzyme digestion followed by phenol/chloroform purification and ethanol precipitation. After electroporation, cells were added to 10 ml of HMI-9 for 24 hour recovery at 37 °C. The recovered cells were centrifuged at 600 x g for 10 minutes at room temperature and resuspended in fresh HMI-9 medium. A volume of  $5 \times 10^6$  cells were diluted in 36 ml of HMI-9 medium containing the appropriate drug concentration and plated over a 24 well plate at 1.5 ml per well, i.e. approximately  $2.08 \times 10^5$  cells per well. Growing cells were counted after 7-10 days by imaging each well of the plate under a light microscope (Leitz). The number of wells that contained transformants was recorded. The population of

transformants in a well were considered clonal, if the number of wells with growing cells did not exceed 10 out of 24.

Once the clones were confirmed, each cell line was stablilated, and stored in liquid nitrogen. This was carried out by adding 0.9 ml of cells at log phase density of  $2 \times 10^6$  per ml with 0.1 ml of 100 % glycerol. This was then gently mixed thoroughly and added to a cyro-vial (Nunc) and frozen at  $-80^\circ\text{C}$  over 48 hours. This was then moved to liquid nitrogen storage and given a Wellcome Centre for Molecular Parasitology (WCMP) stablilate number.

## **2.2 Isolation of DNA**

### **2.2.1 Methods of DNA extraction**

The QIAprep spin miniprep kit was used to extract plasmid DNA from *E. coli* cells. For larger preparations of DNA from *E. coli* cells, GenElute HP plasmid kit (Sigma) was used. The manufacturer's protocols were followed for all extraction methods using Qiagen and Sigma kits.

DNeasy blood and tissue kit (Qiagen) was used to extract DNA from *T. brucei* cells. For larger preparations of DNA from *T. brucei* cells,  $1 \times 10^7$  cells were centrifuged at  $600 \times g$  for 10 minutes and resuspended in 500  $\mu\text{l}$  of lysis buffer [50 mM Tris-HCl (pH 8.0), 1 mM EDTA, 100 mM NaCl], and SDS and Proteinase K (Sigma) were added to final concentrations of 0.01 % and 0.1  $\text{mg.ml}^{-1}$  respectively. This was incubated at  $37^\circ\text{C}$  overnight; protein and lipid were removed using phenol/chloroform extraction and the genomic DNA recovered by ethanol precipitation (see below).

### **2.2.2 Phenol/Chloroform extraction and ethanol precipitation of DNA**

To purify genomic or plasmid DNA from solution, a phenol/chloroform extraction and ethanol precipitation was carried out. This was done by adding an equal volume of buffer-saturated phenol:chloroform:isoamyl alcohol (25:24:1) (Sigma) to the DNA solution. The solution was mixed well by vortexing for plasmid DNA and gently inverted for genomic DNA. The two phases were separated by centrifugation in a microcentrifuge at  $16,000 \times g$  for one minute. The aqueous layer was removed carefully to a new tube.

One tenth of the total volume of 3 M sodium acetate ( $\text{CH}_3\text{COONa}$ ) (pH 5.2) (BDH) was added along with three times the total volume of 100 % ethanol (Fisher). The solution was then frozen at  $-20\text{ }^{\circ}\text{C}$  for 20 minutes, after which the DNA was centrifuged in a microcentrifuge at  $20,000 \times g$  for 30 minutes at  $4\text{ }^{\circ}\text{C}$ . The DNA pellet was washed with  $100\text{ }\mu\text{l}$  of 70 % ethanol and re-centrifuged for one minute. The pellet was air dried and resuspended in the required volume of  $\text{dH}_2\text{O}$ .

## 2.3 RNA extraction

### 2.3.1 RNA extraction from *T. brucei* cells

For extraction of RNA, RNeasy mini kit (Qiagen) was used. The RNA extract was then treated with deoxyribonuclease I (DNase; Invitrogen) to remove any DNA contamination from the sample. The manufacturer's protocols were used for all methods.

### 2.3.2 Reverse Transcriptase-PCR (RT-PCR)

To convert RNA into first-strand cDNA, the SuperScript™ First-Strand Synthesis System for RT-PCR (Invitrogen) was used. The manufacturer's instructions were followed using random hexamers. Briefly,  $5\text{ }\mu\text{g}$  of DNase treated RNA,  $1\text{ }\mu\text{l}$  of random hexamers ( $50\text{ ng}\cdot\mu\text{l}^{-1}$ ), and  $1\text{ }\mu\text{l}$  of 10 mM dNTP mix was added to a final volume of  $10\text{ }\mu\text{l}$  with DEPC-treated  $\text{H}_2\text{O}$ . Two reactions were carried out for each sample, one the positive sample and the other as a negative control containing no reverse transcriptase (RT). Each sample was incubated at  $65\text{ }^{\circ}\text{C}$  for 5 minutes, followed by cooling on ice. To each sample,  $9\text{ }\mu\text{l}$  RNaseOUT reaction [ $2\text{ }\mu\text{l}$  of 10X RT buffer,  $4\text{ }\mu\text{l}$  of 25 mM  $\text{MgCl}_2$ ,  $2\text{ }\mu\text{l}$  of 0.1 M DTT,  $1\text{ }\mu\text{l}$  of RNaseOUT recombinant ribonuclease inhibitor] was added. The reaction was incubated at  $25\text{ }^{\circ}\text{C}$  for 2 minutes. To the positive sample (and not the negative control),  $1\text{ }\mu\text{l}$  (50 U) of SuperScript™ II RT was added and incubated at  $25\text{ }^{\circ}\text{C}$  for 10 minutes. The reaction was then incubated at  $42\text{ }^{\circ}\text{C}$  for 50 minutes followed by  $70\text{ }^{\circ}\text{C}$  for 15 minutes, then chilled on ice. The RNA template was degraded by addition of RNase H to each tube and incubated at  $37\text{ }^{\circ}\text{C}$  for 20 minutes. The cDNA was used as template for amplification using standard PCR (see below).



## 2.4 Cloning procedure

### 2.4.1 Polymerase Chain Reaction (PCR)

PCR reactions were carried out for cloning and screening. Routinely, the PCR mixture contained 1 U of either *Taq* (Biolabs) or Herculanase (Stratagene), 0.5  $\mu$ l (100 pmoles) of the primer pair, see Table 2.1, the required buffer concentration, 1  $\mu$ l (10 mM) dNTPs and 1-5 ng of template DNA to a final volume of 50  $\mu$ l per reaction. DNA from *T. brucei* strain 427 was used as DNA template for cloning of the ORFs and gene fragments. PCR amplification was carried out in either a thermal cycler Robocycler® (Stratagene) or a TC-3000 (Techne), under the PCR conditions: denaturation at 94 °C for 5 minutes, followed by 30 cycles of denaturation at 94 °C for one minute, annealing of primers at 50-60 °C for one minute and extension at 72 °C for one minute per kilobase amplified. A final extension time of 10 minutes at 72 °C ensured the PCR amplification had reached its completion. For purification of the amplified DNA, the PCR cleanup kit (Qiagen) was used.

**Table 2-1: Description of the primers used in the experiments described in this thesis.**

Description of the primers used for PCR amplification in the separate Chapters of this thesis, Chapters 3, 4, and 6. This gives a brief description showing any restriction enzyme sites in colour and gives the specific DNA sequence of the primer. This Table is continued on two following pages, 50-51.

Name	Details	Primer
<b>Chapter 3</b>		
RAD51-4 StartORF-For	Amplification of the Left Hand Flank of <i>RAD51-4</i> <i>XbaI</i>	ccctcttagaagatgacactctggccagt
RAD51-4 StartORF-Rev	<i>EcoRV</i> <i>SpeI</i>	ccccgatatcactagfacatcatccagccgagcgctc
RAD51-4 EndORF-For	Amplification of the Right Hand Flank of <i>RAD51-4</i> <i>EcoRV</i> <i>NcoI</i>	ccccgataccccatgggggagggtaattgaacggca
RAD51-4 EndORF-Rev	<i>XhoI</i>	ccccctcgagagcagtagaacaatgttggc
RECA2 StartORF-For	Amplification of the Left Hand Flank of <i>RAD51-6</i> <i>XbaI</i>	cccccttagaagatggctagcgtgaatgctcc
RECA2 StartORF-Rev	<i>EcoRV</i> <i>SacII</i>	ccccgatatcccgccggcgatcgcatagcgcttctag
RECA2 EndORF-For	Amplification of the Right Hand Flank of <i>RAD51-6</i> <i>EcoRV</i> <i>Apal</i>	ccccgatacgggcccgcgtgggatctctggggcga
RECA2 EndORF-Rev	<i>XhoI</i>	ccccctcgagtcgctgctgtgcataaccttc
Tubulin For <i>EcoRV</i>	Amplification $\beta\alpha$ -blasticidin-actin IR from pCP101 <i>EcoRV</i>	cccgatatctgggtcccatgtttgctc
Actin Rev <i>EcoRV</i>	<i>EcoRV</i>	cccgatatctattttatggcagcaacgag
Tubulin For <i>NruI</i>	Amplification $\beta\alpha$ -puromycin-actin IR from pCP121 <i>NruI</i>	ccctcgcgatgggtcccattgttgcctc
Actin Rev <i>NruI</i>	<i>NruI</i>	ccctcgcgactattttatggcagcaacgag
Upstream of RAD51-4 ORF For	Confirmation of generation of <i>RAD51-4</i> mutants	tggtagaagtgaggagggt
Upstream of RECA2 ORF For	Confirmation of generation of <i>RAD51-6</i> mutants	gtgtttggggagtaggtg
RAD51-4 REX For <i>NruI</i>	Amplification of <i>RAD51-4</i> ORF for re-expressor construct <i>NruI</i>	ccctcgcgaatggatggatgcactctggccag
RAD51-4 REX Rev <i>NruI</i>	<i>NruI</i>	ccctcgcgactagcagtagaacaatgttggc
RECA2 REX-For <i>EcoRV</i>	Amplification of <i>RAD51-6</i> ORF for re-expressor construct <i>EcoRV</i>	cccgataatcaggctagcgtgaatgctc
RECA2 REX-Rev <i>EcoRV</i>	<i>EcoRV</i>	cccgatatctcattcgtcgttgcatataaccttc
51-6 C-term PTP For	Amplification of probe to check <i>RAD51-6</i> re-expressor	gatacggccctatcgcatcgacatgtgtcac
RECA2 Middle Rev	Amplification of probe to check <i>RAD51-6</i> re-expressor	aacaccatcgccctgagc
$\beta\alpha$ IR For	Amplification $\beta\alpha$ IR for the VSG mechanism determination Forward	gtcccattgtttgcctcttc
$\beta\alpha$ IR Rev	Reverse	gcagatagccctcacgcatg

Table 2-1: Description of the primers used in the experiments described in this thesis, continued.

Hygro For	Amplification of hygromycin for VSG mechanism determination	cctcgcgaatgaaaaagcctgaactcacc
Hygro Rev	Reverse	cctcgcgattccttggccctcggacgagt
Neo For	Amplification of neomycin for VSG mechanism determination	cgcattgattgaacaagatgg
Neo Rev	Reverse	tcaagaaggcgatagaaggc
VSG For	Amplification of VSG221 for VSG mechanism determination	atgcctccaatcaggaggc
VSG Rev	Reverse	tgtatcgcgcaactgcag
<b>Name</b>	<b>Details</b>	<b>Primer</b>
<b>Chapter 5</b>		
R51-5 For	Amplification of <i>RAD51-5</i> ( <i>RecA</i> ) ORF <i>EcoRI</i> , <i>Apal</i>	ccccgaattcggggcccatgtctgtgtctccacc
R51-5 Rev	<i>SalI</i>	ccccgtcgcattatcagggtaaaaagatgt
SiteDirMut-RecA For	To make <i>RAD51-5</i> in frame with the addition of <i>A</i> in the <i>Apal</i> site	gatatcagctggggaccatgtctgtgtg
SiteDirMut-RecA Rev	Reverse Complement of the forward primer	cacacagacatgggtccagctgagatc
N-TermR51-4 For	Amplification of the N terminus of <i>RAD51-4 EcoRI</i>	cccgaattcatggatgtggagcttcttac
N-TermR51-4 Rev	<i>PstI</i>	ccccctgcagctagcagtgaacaaatgtg
C-TermR51-4 For	Amplification of the C terminus of <i>RAD51-4 EcoRI</i>	cccgaattcagatgtagcactctggccagt
C-TermR51-4 Rev	<i>PstI</i>	ccccctgcaggccacacttcacgccagcaa
51-5 Cterm PTP F	Amplification of <i>RAD51-5</i> for PTP cloning <i>Apal</i>	gattcggggcccgaggaggactggaggtacagc
51-5 Cterm PTP R	<i>NotI</i>	gattcggggccggcggtaaaagatgttcccacgcg
RAD51-5 For	Amplification of the <i>RAD51-5</i> to check PTP integration	gattcccatggcaacggcgaggagatttcacag
PTP-Cterm Check Rev	Amplification of the <i>RAD51-5</i> to check PTP integration	ggatgaaggcggttctgtg
51-6 Nterm PTP F	Amplification of <i>RAD51-6</i> for PTP cloning <i>NotI</i>	ccggggccgcgcgtagcgtagaatgtccc
51-6 Nterm PTP R	<i>KpnI</i>	ccgggtacctcgtcgtgtcatatccttc
PTP-Nterm Check For	Amplification of the <i>RAD51-6</i> to check PTP integration	gattcagatttagaagatcagggtggatccctcg

Table 2-1: Description of the primers used in the experiments described in this thesis, continued.

Chapter 6		
AB For <i>XbaI</i>	Amplification of the $\alpha\beta$ -IR for luciferase positive control <i>XbaI</i>	cccttagactagaaagtgtgacaaagtc
AB Rev <i>Apal</i>	<i>Apal</i>	ccggggcccggaactattttctttgatgaaag
BA IR For	Amplification $\beta\alpha$ -IR <i>SacI</i>	gagctcgtcccatgtttgtcctcttc
BA IR Rev	<i>EcoRV</i> , <i>XhoI</i> , <i>MluI</i> , <i>SmaI</i> , <i>NcoI</i>	gatactctcgagacggctagccctccatggcgagatagcctcacgcatg
AB IR For	Amplification of $\alpha\beta$ -IR <i>EcoRV</i> , <i>XbaI</i> , <i>Apal</i> , <i>SacII</i>	gatactcttagagggcccccgccgctagaaagtgtgacaacgtc
AB IR Rev	<i>KpnI</i>	gggtaccagccctgaaagcagacgat
Actin For	Amplification of Actin <i>XhoI</i> , <i>NsiI</i> , <i>NotI</i>	ctcgagatgcagcggccgcctaacacccgggtgtgtgg
Actin Rev	<i>EcoRV</i>	gatacgcagcaacgagaccttac
Calmodulin For	Amplification of Calmodulin <i>EcoRV</i>	gatacgcgaattccgtgcctcc
Calmodulin Rev	<i>XbaI</i> , <i>PstI</i>	tctagactgcagatcgatcgtaaatcaagtgg
Ble For	Amplification of Phleomycin <i>EcoRV</i>	cccgatatcatggccaagttgaccagtg
Ble Rev	<i>EcoRV</i>	cccgatatctcagtcctgtcctcctgg
Start Luc For	Amplification of Start-Luciferase Fragment <i>SmaI</i>	cccgagcctgtaaaagccaccatggaag
Start Luc Rev	<i>SmaI</i>	cccgagcctcctcctcggtgtaatcag
End Luc For	Amplification of the End-Luciferase Fragment <i>PstI</i>	ccccgcagtcctcagcggttttga
End Luc Rev	<i>SacII</i>	ccccgcgggtcagtcgtctgtcgaagcg
RecombineCheck For	Amplification to analyse recombination event	cctcgtgagattctgc
RecombineCheck Rev	Reverse	ccgggtatccagatccac
BSD For	Amplification of blasticidin <i>EcoRV</i>	cccgatatcactagtagtggccaagcc
BSD Rev	<i>EcoRV</i>	cccgatcaggatccttagccctcc
End Luc For New	Amplification of the inverted luciferase end <i>SacII</i>	cccccgcggtcccatcagcggttttga
End Luc Rev New	<i>PstI</i> , <i>I-SceI</i>	ccccgtcagtagggataaacagggttaattcatgtctgtcgaagcg

## 2.4.2 Restriction enzyme digestion of DNA

Restriction enzyme digestion of DNA was performed using Biolabs enzymes using the manufacturer's instructions. Routinely, for restriction mapping approximately 10 µg of DNA was digested in a final reaction volume of 20 µl, and for transfecting into *T. brucei* approximately 1 mg of DNA was digested in a final reaction volume of 200 µl.

## 2.4.3 Gel electrophoresis and DNA fragment purification

Standard electrophoresis techniques using 1 % (w/v) agarose LE analytical grade gel (Cambrex) in 1x Tris[hydromethyl]methylamine Acetate buffer (TAE) was employed for routine analysis (Sambrook *et al.*, 1989). Electrophoresis was performed at 100 volts using a Bio-Rad PowerPac Mini. To the gel an aliquot of 1:20,000 dilution of Sybr®Safe stain (Invitrogen) was added before it set. The gel was visualized by a U.V. transilluminator (Bio-Rad) and digitised using Quantity One 4.6 (Bio-Rad).

For extraction of the DNA, fragments were excised from the agarose gel and then extracted using MinElute gel extraction kit (Qiagen) according to the manufacturer's instructions. The DNA was routinely eluted with 30-50 µl TE buffer or dH<sub>2</sub>O.

## 2.4.4 DNA Ligation

DNA ligations were performed using 1:1, 1:3, and 1:0 (as a negative control) ratios of vector:insert in volume, 2.5 U T4 ligase (Biolabs) with 2 µl 10X ligase buffer in a final volume of 20 µl. The reaction took place at 4 °C overnight. Table 2-2 describes the plasmids used in this thesis.

**Table 2-2: Description of the vectors used in the experiments described in this thesis.**  
The names, brief description and source of the vectors used in the experiments in this thesis.

Vector Name	Application	Reference
pCP101	source of $\beta\alpha$ IR- <i>BSD</i> -Actin IR	Proudfoot & McCulloch (2005b)
pCP121	source of $\beta\alpha$ IR- <i>PUR</i> -Actin IR	Proudfoot & McCulloch (2005b)
pRM481	source of $\beta\alpha$ IR- <i>BLE</i> -Actin IR- $\alpha\beta$ IR	McCulloch gift
TOPO 2.1 pCR@2.1	cloning vector	Invitrogen
pBluescript II KS+	cloning vector	Stratagene
pHybLex/Zeo	yeast-two hybrid "bait" vector	Invitrogen
pYesTrp2	yeast-two hybrid "prey" vector	Invitrogen
pC-PTP	PTP tagging the C terminus of a protein	Schimanski <i>et al.</i> (2005b)
pN-PTP	PTP tagging the N terminus of a protein	Schimanski <i>et al.</i> (2005b)
pGL3-Basic	source of <i>Luciferase</i> ORF	Promega

## 2.4.5 Transformation into *E. coli*

All *E. coli* growth was on Luria broth [LB; 5 g yeast extract (Melford), 10 g tryptone (Formedium), 5 g NaCl (low salt) or 10 g NaCl (high salt) in 1 L pH 7.0] and agar (Melford) was added if plates were required. Ligations were transformed into chemically competent *E. coli* XL1-Blue MRF [ $\Delta(mcrA)183$   $\Delta(mcrCB-hsdSMR-mrr)173$  *endA1 supE44 thi-1 recA1 gyrA96 relA1 lac* [F' *proAB lacI<sup>q</sup>ZAM15*]; Stratagene] by heat shocking the cells at 42 °C for 45 seconds with the required amount of DNA. An aliquot of 900  $\mu$ l of LB was added to the cells and cells were incubated for 50 minutes at 37 °C. Approximately 200  $\mu$ l cells were then plated onto selective LB agar plates containing 40  $\mu$ g.ml<sup>-1</sup> of 5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactoside (X-gal; Sigma) and 100  $\mu$ g.ml<sup>-1</sup> of ampicillin (Sigma). The plates were incubated at 37 °C overnight. Addition of X-gal allowed blue/white selection; a few white colonies were screened by restriction digestion and positives were sent to be sequenced by the University of Dundee DNA sequencing service. The M13 forward and reverse primers were usually used for sequencing. All alignments and manipulations were carried out using BlastN (Pubmed), AlignX (Invitrogen) and Vector NTI Suite 9 (Invitrogen).

## 2.4.6 Transformation into *S. cerevisiae*

Preparation of competent *S. cerevisiae* L40 strain [MATa *his3 $\Delta$ 200 trp1-901 leu2-3 122ade2 LYS2::(4 lexAop-HIS3) URA3::(8lexAop-lacZ) GAL4*; Invitrogen] was performed using the *S. c.* EasyComp™ Transformation Kit (Invitrogen) following manufacturer's instructions. Co-transformations of both pHybLex/Zeo-Bait and pYesTrp2-Prey (Invitrogen) into yeast L40 strain was carried out and selected using minimal media (see

Section 2.10 for media details). The pHybLex/Zeo-Bait and pYesTrp2-Prey constructs were made by Sandra Terry.

## 2.4.7 TOPO TA cloning

For one-step cloning of a PCR fragment into a plasmid vector, Invitrogen's TA cloning kit was used. This involved direct insertion of *Taq* polymerase-amplified PCR products into the plasmid vector pCR®2.1-TOPO®. A reaction of 1.0 µl of PCR product (approximately 10 ng), 1 µl of NaCl solution, 1 µl pCR®2.1 vector, and 3 µl dH<sub>2</sub>O was incubated at room temperature for 5 minutes. Then, 2.0 µl of this reaction was added to the OneShot® competent cells (Invitrogen). The cells were transformed by heat shocking (see Section 2.4.5).

## 2.5 Southern blotting

Southern blotting analysis (Southern, 1975) demonstrated that the alleles of *RAD51-4* and *RAD51-6* were disrupted. The purified genomic DNAs from wildtype, heterozygous, and homozygous mutants were digested using restriction enzymes at 37 °C overnight. These were then electrophoresed on a 0.8 % agarose gel overnight as previously described (see Section 2.4.3). The gel was imaged on a U.V. transilluminator (Bio-Rad) with a ruler parallel to the gel so that the sizes of the bands detected by hybridisation could be measured. To nick the DNA, the gel was soaked in 0.25 M HCl (Fisher) for 15 minutes and washed in dH<sub>2</sub>O. The DNA was then denatured by soaking the gel in denaturing solution [0.5 M NaOH and 1.5 M HCl (Sigma)] for 30 minutes. This denaturing solution was washed off with dH<sub>2</sub>O and the gel was soaked in neutralising solution [1 M Tris-HCl pH 8.0, 1.5 M NaCl (Sigma)] for a further 30 minutes. The DNA was transferred to a nylon membrane (Hybond-XL; Amersham) by capillary blotting using 20 X SSC [3 M NaCl, 0.3 M Na<sub>3</sub>C<sub>6</sub>H<sub>5</sub>O<sub>7</sub>·2H<sub>2</sub>O] as a transfer buffer overnight. The DNA was crosslinked to the membrane using a UV stratalinker (Stratagene).

PCR-amplification (see Section 2.4.1) provided the DNA fragments used for the probe. These products were gel extracted (see Section 2.4.3). The probe was made using the Gene Images Random Prime Labelling Kit from Amersham and hybridisation of the probe to the membrane was performed following the manufacturer's instructions. The probe was detected using the *Gene Images CDP-Star* Detection Module (Amersham).

## 2.6 Protein analysis

### 2.6.1 Protein extraction from *S. cerevisiae* cells

The following protocol, adapted from procedures described by Hybrid Hunter™ Invitrogen, was used to extract proteins from yeast co-transformants. A single colony transformant was inoculated into 10 ml of the appropriate media and grown at 30 °C while shaking at 250 x *g* for 24 hours. The cells were pelleted by centrifugation for 5 minutes at 2500 x *g* and the supernatant was decanted. The cell pellet was frozen for 10 minutes at – 80 °C, then thawed and resuspended in 100 µl of pre-warmed cracking buffer [8 M urea (BDH), 5 % SDS (BDH), 40 mM Tris-HCl pH 6.8 (Fisher), 0.1 M EDTA, 1 % β-mercaptoethanol, 0.4 mg.ml<sup>-1</sup> bromophenol blue (Sigma)]. The cell suspension was transferred to a tube containing 100 µl of glass beads (Sigma) and then incubated at 70 °C for 10 minutes. The solution was mixed by vortexing for one minute and centrifuged at 14,000 x *g* for 5 minutes at room temperature. 50 µl of the supernatant was boiled at 100 °C for 5 minutes and electrophoresed on a 10 % polyacrylamide SDS gel (Bio-Rad).

### 2.6.2 Protein extraction from *T. brucei* cells

#### 2.6.2.1 *Crude protein extraction from T. brucei cells*

For protein preparations from *T. brucei* cells, 2 x 10<sup>8</sup> cells were pelleted at 1000 x *g* for 10 minutes and washed twice in 1 ml of PBS. The cells were re-centrifuged and resuspended in 200 µl lysis buffer [500 mM Tris HCl (pH 6.8), 10 % glycerol, 10 % SDS, 5 % β-mercaptoethanol, 0.05 % bromophenol blue (Sigma)]. The extract (20-50 µl) was electrophoresed on a 10 % polyacrylamide SDS gel (Bio-Rad).

#### 2.6.2.2 *Whole cell and nuclear protein extraction from T. brucei cells*

For larger protein preparations from *T. brucei* cells, 1 x 10<sup>10</sup> cells were centrifuged at 1000 x *g* for 10 minutes and washed twice in 1 ml of PBS. The cell pellet was frozen on dry ice and stored at -80 °C until required, for up to three months. The protocol was adapted from Smeaton *et al.* (2007). Frozen pellets were resuspended in 300 µl of hypotonic lysis buffer [HLB: 10 mM Tris-HCl (pH 7.9), 1 mM EDTA, 10 mM NaCl, 1 mM DL-dithiothreitol (DTT) (Sigma) and protease inhibitor cocktail (Roche)], incubated on ice for 20 minutes and vortexed for 30 seconds. The mixture was centrifuged at 800 x *g* for 2 minutes at room temperature, to collect the nuclei. The supernatant was retained. The pelleted nuclei were



resuspended in nuclear extract buffer [NEB: 20 mM Tris-HCl (pH 7.9), 1 mM EDTA, 400 mM NaCl (Sigma), 1 mM DTT, protease inhibitor cocktail (Roche), and 10 % glycerol (Riedel deHaën)] and incubated on ice for 20 minutes. The retained supernatant was added back to the reaction and the cell debris was removed by a centrifugation at 20,000 x g for 20 minutes at 4 °C. The supernatant was aliquoted and quickly frozen using dry ice and stored at -80 °C (Smeaton *et al.*, 2007). The protein concentration of the extract was quantified using the Bradford assay (Bio-Rad).

### **2.6.3 Electrophoresis of SDS-polyacrylamide gels and Western blotting**

To examine expression of protein in *T. brucei*, cell lysates were analysed by western blot (Towbin & Gordon, 1984). Protein samples were loaded onto a 10 % polyacrylamide gel (Bio-Rad ready gels) and electrophoresed at 180 V in 1 X SDS running buffer (Bio-Rad) for approximately one hour or until required separation was achieved.

For western blotting, the proteins were then transferred to a nylon membrane (Amersham), at 100 V for 60-70 minutes in transfer buffer [25 mM Tris, 192 mM glycine and 20 % methanol (Sigma)]. The membrane was first blocked for one hour then probed with the appropriate concentration of primary antibody in blocking solution [5% powdered-milk (Marvel), 0.01% tween20 (Sigma) in PBS] for 1-18 hours. The membrane was washed with blocking solution and probed with the appropriate concentration of secondary antibody for one hour at room temperature. To detect antibody binding, horseradish peroxidase (HRP) fused secondary antibodies were used. The HRP was detected by enhanced chemi-luminescence (Supersignal by Pierce). Film was exposed to the chemi-luminescence. This film was processed in developer and fixer using a Kodak Ex-omat. Most membranes were then stripped with stripping buffer (Pierce) and re-probed with a different antibody.

## **2.7 Purification of anti-RAD51 and the anti-RAD51 paralogue anti-sera**

Antibodies were raised against the RAD51 and RAD51 paralogue purified proteins containing a C-terminal 10 X HIS tag (RAD51-HIS and RAD51 paralogues-HIS); this was carried out by Dr Chris Stockdale. One rabbit was used to raise anti-serum against the RAD51 and RAD51-4 recombinant proteins, a sheep was used to raise anti-serum against RAD51-3, two chickens were used to raise anti-serum against RAD51-5, and a rat was

used to raise anti-serum against RAD51-6 by the Scottish Blood Transfusion Board, Edinburgh.

### **2.7.1 Anti-RAD51-5 purification from chicken egg yolks**

Pierce® Chicken IgY Purification Kit was used to purify IgY from the egg yolk of chickens. This was carried out using the manufacturer's instructions. Briefly, the total volume of the egg yolk was calculated. Five times this volume of delipidation reagent was added to the yolk. This mixture was mixed thoroughly using a magnetic stirrer overnight at 4 °C. Then the mixture was pelleted at 10,000 x g for 15 minutes at 4 °C. The supernatant volume was measured and filtered to remove all particulate material. An equal volume to the supernatant of IgY precipitation reagent was added and mixed for two minutes. This mixture was left overnight at 4 °C. After this settling period, the mixture was gently mixed and then centrifuged at 10,000 x g for 15 minutes at 4 °C and the supernatant was discarded. The pellet obtained contains the purified IgY. To dissolve the IgY out of the pellet, an equal amount to the original egg yolk volume of PBS was added and gently mixed until it was completely dissolved.

### **2.7.2 Affinity Purification of anti-RAD51 and the anti-RAD51 paralogue anti-sera**

For purification of the anti-sera, purified proteins containing an N-terminal GST tag were used (GST-RAD51 and GST-RAD51 paralogue proteins). These GST protein purifications were coupled to beads (Aminolink resin; Pierce) which were then used as an affinity column. The protein purifications and affinity columns were prepared by Dr Chris Stockdale. Once made, the affinity column was stored in 0.05 % sodium azide/PBS at 4 °C.

For affinity purification of the anti-sera, the affinity column was allowed to rest at room temperature. The resin was washed with 5 ml of 0.05 % sodium azide/PBS (Sigma). The anti-sera was added to the column in 750 µl aliquots and allowed to flow through. This was repeated until the entire sample had been loaded to the column. 100 µl of 0.05 % sodium azide/PBS was added to the column and allowed to flow through, after which the column was closed at the bottom. 500 µl of 0.05 % sodium azide/PBS was added to the column and the top was capped. This was incubated at room temperature upright for one hour, after which, the column was washed with 12 ml of PBS. The antibody was eluted off the column by addition of 4 ml of 100 mM glycine-HCl (pH 2.7). The elutant was collected in 0.5 ml fractions, and quickly neutralized with the addition of 50 µl of 1 M Tris-HCl (pH 7.5). The

column was washed with 10 ml of PBS and 2 ml of 0.05 % sodium azide/PBS was added as a storage buffer.

The protein concentrations of the fractions were analysed using the Bradford assay (Bio-Rad). As the Bio-Rad assay could only tolerate 2.5 mM of glycine, the fractions were diluted to a ratio of 1:40 with dH<sub>2</sub>O. The fractions with the highest concentrations of protein were pooled and their buffer was exchanged into PBS using the Bio-Rad Econo-Pac 10DG column, using the manufacturer's instructions.

### **2.7.3 Acetone protein extraction for purification of anti-RAD51-4 anti-serum**

The acetone protein extraction method was adapted to purify anti-serum, in this case, anti-RAD51-4 anti-serum (Sambrook *et al.*, 1989). This method uses an acetone extract of a cell line, in this case *T. brucei rad51-4* *-/-*, that is known not to express the target antigen, in this case *T. brucei* RAD51-4. The cells were grown in 500 ml to a density of approximately  $2 \times 10^6$  per ml, and were pelleted, washed in PBS and finally resuspended in 8 ml of PBS and left on ice for 10 minutes. After chilling, 64 ml of cold acetone was added to each and left on ice for a further 30 minutes. This solution was pelleted at  $11,600 \times g$  for 10 minutes and the supernatant discarded. 10 ml of cold acetone was added to each and left on ice for a further 10 minutes. The solution was then centrifuged again at  $11,600 \times g$  for 10 minutes and the supernatant discarded. The pellet was transferred to filter paper (Wattman) and spread out as much as possible to allow the mixture to completely dry. This dried protein was then weighed. For every 1 gram of dried protein, 100 ml of unpurified anti-serum can be added. In this case, 1 ml of anti-RAD51-4 anti-serum was added in a 1 ml eppendorf along with the dried protein/cell mixture. This was then incubated at 4 °C for 30 minutes and then pelleted at  $11,600 \times g$  for 30 minutes. The supernatant contains the purified antibodies.

## **2.8 Analysing the knockout phenotypes**

### **2.8.1 Growth curve and doubling times**

To determine if the growth of *rad51-4* *-/-* and *rad51-6* *-/-* mutants differed to wildtype cells, population doubling time was measured. To examine this, wildtype, heterozygous and homozygous mutants were inoculated (in triplicate) at  $5 \times 10^4$  cells per ml in HMI-9 and cell densities were counted at fixed time points, at 24 hour intervals. The results were

plotted on a semi-logarithmic graph (MS Excel, Microsoft). The doubling time for the cell lines were calculated by comparing the cell count in mid-log phase using the following equation:  $\text{Time Elapsed} \div (3.3 \times \log(\text{Cell count at the end of time elapsed} \div \text{Cell count at the beginning}))$ , where the time elapsed is in hours, and the cell count is in cells per ml.

## **2.8.2 DNA damage sensitivity**

### **2.8.2.1 *Clonal survival with DNA damaging agents***

The cell lines were grown to mid-log phase, approximately  $2 \times 10^6$  cells per ml and the cells were counted using a haemocytometer. The cells were then diluted to 1 cell per 200  $\mu\text{l}$  of HMI-9 medium. The diluted cells were plated over a 96 well plate with 200  $\mu\text{l}$  in each well (i.e. 1 cell per well). Five different concentrations of methyl methanesulphonate (MMS) and phleomycin were added to the cells at final concentrations of 0.0 %, 0.0001 %, 0.0002 %, 0.0003 %, 0.0004 %, and 0.0  $\mu\text{g}.\text{ml}^{-1}$ , 0.025  $\mu\text{g}.\text{ml}^{-1}$ , 0.05  $\mu\text{g}.\text{ml}^{-1}$ , 0.075  $\mu\text{g}.\text{ml}^{-1}$ , 0.1  $\mu\text{g}.\text{ml}^{-1}$ , respectively. These plates were left in the incubator for 14 days. Growth per well was checked and at day 14. This was done in triplicate and the average clonal survival was calculated.

### **2.8.2.2 *Alamar blue analysis of $IC_{50}$ values for MMS and phleomycin***

The cell lines were grown to mid-log phase, approximately  $2 \times 10^6$  cells per ml and the cells were counted using a haemocytometer. First, the 96 well plate (Nunc) was prepared; 100  $\mu\text{l}$  of HMI-9 media was added to all the wells except for the first column. To the first column, 200  $\mu\text{l}$  of the stock of MMS and phleomycin were added (0.01 % and 4  $\mu\text{g}.\text{ml}^{-1}$  respectively). Serial dilutions were made, by taking 100  $\mu\text{l}$  from the first column and mixing it with the next column. With clean pipette tips, 100  $\mu\text{l}$  from the second column was added to the third column and mixed, and so on. To each well, 100  $\mu\text{l}$  of cells at a density of  $2 \times 10^5$  cells per ml were added. These cells were allowed to grow for 48 hours after which 20  $\mu\text{l}$  of 0.125 % resazurin/PBS (Sigma) was added. This was then incubated for 24 hours at 37 °C. The plate was read by a spectrometer using filter of 540 nm excitation and 590 nm emission (Wallac Envision, 2102 Multi-label reader) and the data was analysed using MS Excel and GraphPrism.

## **2.8.3 Transformation efficiency assay**

To assay for recombination rate, a linearised piece of DNA containing a drug resistant cassette was transformed into each cell line. A maxi preparation (GenElute HP plasmid kit,

Sigma) of pRM450 was digested for which approximately 200 µg of cleaned linearised plasmid was required. This construct contained a phleomycin resistance cassette flanked with tubulin intergenic regions. The cell lines were grown to mid-log phase, approximately  $2 \times 10^6$  cells per ml, in a final volume of 25 ml. The cell density was measured by counting with a haemocytometer. A total of  $5 \times 10^7$  cells were transfected with 5 µg linearised DNA (see Section 2.1.2). The cells were allowed to recover for three generations before plating out on selective HMI-9 medium containing  $2 \mu\text{g.ml}^{-1}$  of phleomycin. A total of  $5 \times 10^6$  of wildtype and heterozygous mutant cells were diluted in 36 ml of selective media. Then 1.5 ml of the diluted cells were plated in each well of a 24 well plate, which is equivalent to  $2.08 \times 10^5$  cells per well. As the homozygous mutant cells were more likely to have a deficiency in recombination ability,  $2 \times 10^7$  cells were plated over 48 wells, which is equivalent to  $4.16 \times 10^5$  cells per well. The cells were incubated for 14 days at 37 °C and the number of wells with cell growth was scored. Each cell line was analysed in triplicate with the same linearised DNA, for consistency. The transformation efficiency was described as the number of wells with growth per  $10^6$  cells plated out on selective medium.

## 2.8.4 VSG switching frequency and mechanism analysis

Use of the *T. brucei* BSF strain 3174.2 enabled the analysis of the VSG switching frequency and mechanisms (McCulloch *et al.*, 1997; McCulloch & Barry, 1999; Rudenko *et al.*, 1996). This was carried out by generating acquired immunity against *T. brucei* cells expressing a specific VSG in mice, and then re-infecting the mice with *T. brucei* cells that were allowed to switch their VSG *in vitro*. The *T. brucei* cells were recovered from the mice, only those that switched to a different VSG could have survived the host immune system. Female adult ICR mice were used and each was approximately 20 grams in weight. Analysis of the switched *T. brucei* cells recovered from the mice allowed VSG switching frequency and mechanisms to be calculated, this is described in detail in Chapter 3, Section 3.4.5, whilst the experimental details are described below. All animal handling was carried out by Dr Richard McCulloch.

To generate mice immune to a specific VSG the following method was used. Wildtype cells strain 3174.2 was engineered to express a specific VSG, named VSG221, by selecting the cells on HMI-9 containing hygromycin B ( $5.0 \mu\text{g.ml}^{-1}$ ; Calbiochem) and G418 ( $2.5 \mu\text{g.ml}^{-1}$ ; Calbiochem) for a period of 5 days (McCulloch *et al.*, 1997; McCulloch & Barry, 1999; Rudenko *et al.*, 1996). A total of  $2 \times 10^5$  cells expressing VSG221 were injected intraperitoneally into each mouse. For each cell line to be examined, at least three individual mice were injected. When the infection reached  $3\text{-}7 \times 10^8$  cells per ml (typically

3-4 days), the mice were cured by injection of cymelarsan (Rhone Merieux; 5 mg.kg<sup>-1</sup>). The cured mice were then allowed to recover for at least 1-5 weeks prior to further challenge.

To generate switched variants *in vitro*, each cell line which had been previously grown on hygromycin B (5.0 µg.ml<sup>-1</sup>) and G418 (2.5 µg.ml<sup>-1</sup>) in HMI-9 for a period of 5 days to express VSG221, were allowed to grow for a period of nine generations off antibiotic selection. This was carried out by passaging the cells into 25 ml of HMI-9 at a specific dilution such that the final concentration was approximately 2 x10<sup>6</sup> cells per ml: for example, a volume of 7.8 x 10<sup>3</sup> wildtype cells were added to 25 ml of HMI-9 and allowed to grow for 63 hours without passage. Three independent passages were set up of each cell line, and the differences in *in vitro* growth rate were also accounted for.

Once the cells had grown for nine generations, each culture was counted and a total of 4 x 10<sup>7</sup> cells were pelleted at 600 x g and resuspended in 200 µl of HMI-9 media. This suspension was then used to challenge the mice. After a period of 24 hours post-challenge, the surviving *T. brucei* cells were recovered from the mice by cardiac puncture. The *in vivo* growth rates were not examined for the experiments described in this thesis, as there was only a modest defect in *in vivo rad51* <sup>-/-</sup> mutant population doubling times compared to that of the wildtype, as described by McCulloch & Barry (1999). An aliquot of 400 µl of the recovered blood was transferred into a sterile Greiner tube, and centrifuged at 4500 x g for 5 minutes. This separated the blood into different layers, serum, buffy coat (containing white blood cells, platelets and *T. brucei* cells) and red blood cells. Since there was sufficient blood recovered from each mouse, this was done in duplicate. The number of *T. brucei* cells recovered from the mouse could not be counted because the switch frequency was so low, as only eight wildtype 3174.2 cells switch and survive per 10<sup>7</sup> cells injected. To ensure the maximum number of *T. brucei* cells were collected, the serum, buffy coat and the top of the red blood cell layer were transferred to 20 ml of HMI-9 medium. This was then mixed well and distributed over a 96 well plate for each duplicate pair. These plates were incubated at 37 °C for 14-21 days. The growing populations were microscopically monitored and 10 clones from each mouse were passaged into HMI-9 medium. To analyse the mechanism of VSG switching, the DNA was extracted from these cells as previously described (see Section 2.2.1) and replica passages of each selected clone were carried out in media containing hygromycin B (5.0 µg.ml<sup>-1</sup>) or G418 (2.5 µg.ml<sup>-1</sup>).

The switching frequency was calculated by counting the number of recovered *T. brucei* cells over 96 well plates. This number was multiplied by 2.5, as the recovered blood from the mouse was only 40 % of the total approximate blood volume of the animal. This number was then divided by the number of doubling times that had occurred during the 24 hour period following injection into the mouse. A selection of recovered clonal populations were plated on hygromycin or G418, and the DNA from each population was extracted. PCR-amplification and drug sensitivity allowed the VSG switching mechanism to be deduced and the three possible switching events are described in Chapter 3, Section 3.4.5, depending of the presence of the ORF of hygromycin phosphotransferase, neomycin phosphotransferase, and *VSG221* genes, as well as analysing whether or not these are expressed by selection with hygromycin B ( $5.0 \mu\text{g}.\text{ml}^{-1}$ ) and G418 ( $2.5 \mu\text{g}.\text{ml}^{-1}$ ).

## 2.9 Immuno-localisation

### 2.9.1 Fixing *T. brucei* cells to slides

Cells were grown to a density of approximately  $1 \times 10^6$  cells per ml, and pelleted at  $600 \times g$  for 1 minute. The supernatant was removed, and the cells were completely resuspended in 100  $\mu\text{l}$  PBS (Sigma). To the evenly resuspended cells, 1 ml of 1 % (v/v) formaldehyde/PBS was added and the solution was inverted several times. This was then stored at  $4^\circ\text{C}$  for 1-24 hours. The formaldehyde fixed cells were then centrifuged at  $600 \times g$  for 1 minute at  $4^\circ\text{C}$ . The pellet was washed twice in 1 ml chilled PBS followed by a wash of 1 % Triton X 100 (Sigma) to a final concentration of 0.1 %. This was incubated at room temperature for 10 minutes and glycine (Sigma) was added to a final concentration of 0.1 M. Another incubation at room temperature commenced for 10 minutes. The cells were pelleted and resuspended in 200  $\mu\text{l}$  of PBS and spread evenly onto a glass slide (Menzel-glaser). The cell suspension was allowed to dry completely for 2-18 hours.

### 2.9.2 *in situ* fluorescence microscopy and DAPI staining of *T. brucei* cells

Once the cells had been fixed and dried on to the slide (Section 2.9.1), the cells were ready to probe with indirect fluorescent antibodies (IFA) and counter stain with 4', 6-diamidino-2-phenylindole (DAPI; VectaShield mounting medium with DAPI, Vector Labs). First the slides were washed with 1 x PBS for 5 minutes in Coplin jars. The cells were then blocked with TB buffer (0.1 % Triton X 100, 0.1 % BSA in PBS) for 15 minutes. This was removed and the primary antibody which was diluted in TB buffer was added and left for

55 minutes. The primary antibody was washed off by soaking the slides thrice, in PBS for 5 minutes each time. The secondary antibody diluted in TB buffer was added for 45 minutes. The slides were covered to prevent light bleaching of the sample. The slides were then washed as before and counter stained with two drops of VectorShield (Vector Labs). A cover slip was then placed on the slides and sealed with nail polish (Boots). Once dried, the slides were visualised using an Axioskop 2 microscope and Openlab 5.00 software. The images were cropped using PhotoShop 6 (Adobe).

## 2.10 Assaying yeast two-hybrid interaction

*S. cerevisiae* wildtype L40 strain was grown on YPD media (10 g of yeast extract (Melford), 20 g of bactopectone (Melford), and 20 g of dextrose (Fisher) in 1 L at pH 5.8) at 30 °C for 24 hours. If plates were required, agar was added per manufacturer's instructions (Melford). The L40 strain also requires adenine (0.01 %; Sigma) for growth. Minimal media (YC-W; Clontech) was used for the growth of cells containing pYesTrp2-Prey constructs, while selection with 300 µg.ml<sup>-1</sup> zeocin (Invitrogen) was used for cells containing pHybLex/Zeo-Bait constructs. Cells co-transformed with both pHybLex/Zeo and pYesTrp2-Prey were grown up on YC-W media with 300 µg.ml<sup>-1</sup> zeocin at 30 °C for at least 48 hours. The pHybLex/Zeo-Bait and pYesTrp2-Prey plasmids are described in more detail in Chapter 5, Section 5.3.

### 2.10.1 To test for histidine auxotrophy

This was performed by inoculating 10 µl of yeast overnight culture onto an agar plate containing minimal media lacking histidine and containing increasing concentrations of 3' aminotriazole which is a competitive inhibitor of imidazole-glycerolphosphate dehydratase, the product of the *HIS3* gene (Kanazawa *et al.*, 1988). The same yeast overnight cultures were used to inoculate medium containing histidine as a positive control. The cultures from this plate were then used in the β-galactosidase filter lift assay (see Section 2.10.2).

### 2.10.2 β-galactosidase filter lift assay

A colony of each pHybLex/Zeo-Bait and pYesTrp2-Prey co-transformation, X and Y, was grown up on selective media, on YC-W with 300 µg.ml<sup>-1</sup> zeocin at 30 °C for 48 hours. The colonies were transferred to a nitrocellulose membrane (Hybond-C super, Amersham). The filter was cooled on a foiled boat over liquid nitrogen for 30 seconds then submerged for 5 seconds to permeabilise the cells. The filter was thawed and placed in the lid of a Petri dish



containing a filter paper (#1 Whatmann) saturated with 1.5 ml of Z buffer (60 mM  $\text{Na}_2\text{HPO}_4$ , 40 mM  $\text{NaH}_2\text{PO}_4$ , 10 mM KCl, 1 mM  $\text{MgSO}_4$ , (Sigma) pH 7.0) containing 30  $\mu\text{l}$  of 50  $\text{mg}\cdot\text{ml}^{-1}$  of X-gal (Sigma). This was wrapped in moist tissue and incubated at 30 °C for 30 minutes. The blue colouration of the colonies showed the  $\beta$ -galactosidase activity occurred, when the X-gal was cleaved forming 5-bromo-4-chloro-3-hydroxyindole. This by-product was then oxidised, resulting in an insoluble blue product, 5,5'-dibromo-4, 4'-dichloro-indigo. The plate was then allowed to cool to 4 °C and a photograph was taken of it.

### 2.10.3 Yeast $\beta$ -galactosidase assay kit

The yeast  $\beta$ -galactosidase assay kit by Pierce was used to detect and quantify transcriptional activation of  $\beta$ -galactosidase gene, following manufacturer's instructions. Briefly, two independent clones of each pHybLex/Zeo-Bait and pYesTrp2-Prey co-transformant were grown as liquid cultures in YC-W with 300  $\mu\text{g}\cdot\text{ml}^{-1}$  zeocin at 30 °C for 48 hours. Cell density was quantified by measuring the optical density (OD) at 620 nm. Cell lysis solution and o-nitrophenyl- $\beta$ -D-galactopyranoside (ONPG) was added to the cells.  $\beta$ -galactosidase mediated conversion of ONPG to o-nitrophenol, a yellow precipitate, was quantified by measuring absorbance at 420 nm using a 96 well plate reader (Wallac Envision, 2102 Multi-label reader). Each measurement was performed in triplicate.  $\beta$ -galactosidase activity was calculated using the following equation:  $(1000 \times \text{Absorbance } 420 \text{ nm}) \div (t \times V \times \text{OD } 620\text{nm})$  where t is the assay reaction time in minutes and V is the volume of culture used in the assay in ml. The calculation and graphs were completed in MS Excel (Microsoft).

## 2.11 Co-immunoprecipitation

The Pierce ProFound™ co-immunoprecipitation kit was used to isolate native protein complexes from *T. brucei* extract. The whole cell protein extraction from *T. brucei* method was used (see Section 2.6.2.2); this allowed extraction a large concentration of proteins without the presence of cell debris. Pierce's ProFound™ co-immunoprecipitation protocol was followed. Briefly, the antibody coupling gel (AminoLink® Plus Gel) was washed three times with 400  $\mu\text{l}$  of Coupling Buffer (BupH™ Modified Dulbecco's PBS). Approximately 400-500  $\mu\text{g}$  of concentrated affinity purified antibody (see Section 2.7) was coupled to the antibody coupling gel. An overnight incubation at room temperature was performed to allow the resin to couple to the antibody, after which approximately 1-2 mg of total extracted protein was added to the antibody-resin mixture.

## 2.12 PTP purification of protein complexes

A standard purification was carried out with 4 ml of protein extracted using the method described above, Section 2.6.2.2. Firstly 200  $\mu$ l of the IgG beads were equilibrated in a poly-prep chromatography column (BioRad) with 50 ml of PA-150 [150 mM KCl, 20 mM Tris-HCl pH 7.7, 3 mM  $MgCl_2$ , 0.5 mM DTT and 0.1 % of Tween 20 (Sigma)]. The extracted proteins were pooled and 0.5 ml of protease inhibitor tablet (Complete Mini, Roche) dissolved in 1 ml of PA-150 was added. Tween20 was then added to a final concentration of 0.1 % (v/v) to the extract. The column was closed at the bottom and the cell extract was added to the IgG matrix. The top was then closed and left to rotate at 4 °C for two hours at a low speed. After this period of time, the column was then placed in the upright position on a clamp and beads were left to settle. The flow through was then collected, and washed with 10 ml of PA-150 by rotation for 10 minutes at 4 °C. The matrix was washed again with another 15 ml of PA-150 by flow through. All flow through washes were kept.

For further purification, TEV protease removes the Protein A and IgG. This was carried out by washing the IgG-protein A-bound proteins with 15 ml of TEV buffer [150 mM KCl, 20 mM Tris-HCl pH 7.7, 3 mM  $MgCl_2$ , 0.5 mM EDTA, 1.0 mM DTT and 0.1 % of Tween 20 (Sigma)] by rotation for 10 minutes at 4 °C. 2 ml of the TEV buffer and 200 Units of TEV protease was added to the beads and left to rotate in the column for 16 hours at 4 °C.

The anti-protein C affinity matrix was allowed to equilibrate in a second poly-prep chromatography column (BioRad) with 50 ml of PC-150 [150 mM KCl, 20 mM Tris-HCl pH 7.7, 3 mM  $MgCl_2$ , 1 mM  $CaCl_2$ , and 0.1 % of Tween 20 (Sigma)]. The top and bottom plugs from the IgG column were removed and the eluate was recovered by gravity flow and the columns dead volume was eluted by addition of 4 ml of PC-150. To titrate the EDTA in the TEV buffer, 7.5  $\mu$ l of 1 M  $CaCl_2$  was added as well as 0.5 ml of the dissolved protease inhibitor tablet (Roche). This was then transferred to the anti-protein C affinity matrix equilibrated column. The column was closed at the bottom and top plugs were closed. This was allowed to rotate at 4 °C for two hours at a low speed. After this period of time, the column was then placed in the upright position on a clamp and the beads were left to settle. The flow through was then collected. The matrix was washed with a total of 60 ml of PC-150 buffer, with the first and last washes rotating. When the washes were complete, the bottom plug was closed and 600  $\mu$ l of EDTA elution buffer [5 mM Tris-HCl pH 7.7, 10.0 mM EGTA, 5.0 mM EDTA] was added to the matrix. This was incubated for

10 minutes at room temperature and the eluate was collected. This was repeated three times in total.

The eluted protein was then concentrated using the StrataClean resin (Stratagene) by incubating the 10  $\mu$ l of the resin with the eluate for 20 minutes rotating at room temperature. The beads were then pelleted by centrifugation at 5,000 x *g* for 1 minute. The supernatant was removed and the proteins were released from the beads by boiling for 5 minutes in 20  $\mu$ l of SDS-PAGE loading buffer. This was then loaded on to a 10 % SDS-PAGE gel.

## 2.13 Bioluminescence detection

*T. brucei* bloodstream form cells expressing the *luciferase* ORF were analysed to confirm that there were detectable levels of bioluminescence. The cells were counted and washed in PBS to remove the HMI-9, and serial dilution of the cells was carried out in PBS. The cells were diluted from a starting density of  $2 \times 10^6$  to a final density of  $1.56 \times 10^4$  cells in a total volume of 200  $\mu$ l of PBS. To each well, 20  $\mu$ l of 1 mM of d-luciferin/PBS (Promega) was then added. The plate was incubated at 37 °C for 30 minutes. After this, the plate was removed from the incubator and left at room temperature for 15 minutes. Bioluminescence was then read using a 404 nm filter (Envision). Bioluminescence was measured in Relative Light Units and a graph was generated using MS Excel.

**Chapter 3. Role of two *RAD51*  
paralogues in *T. brucei*, *RAD51-4*  
and *RAD51-6*.**

### 3.1 Introduction

DNA sequencing of the genome of *T. brucei* has determined that there are 9,068 predicted genes within the 11 megabase chromosomes (Berriman *et al.*, 2005). Out of these predicted genes only 50.5 % have been assigned a function. Reverse genetics approaches allow gene function to be ascertained, based on the ability to predict function from the protein sequence. These approaches identify a gene from the DNA sequence, disrupt or delete the main open reading frame (ORF) or products of the ORF, and analyse the resulting phenotype. Gene disruption techniques, such as generating null mutants, can help understand the gene product's effect on a phenotype, and allow the biological role of the protein to be established. The complete genome sequence and reverse genetic approaches at the disposal of the researcher are valuable tools and allow pin pointed genotyping.

One such prediction was made by Proudfoot & McCulloch (2005a), where six single copy *T. brucei* *RAD51*-related genes were identified. One gene encoded *RAD51*, while another encoded the closely-related factor, *DMC1*, which is inferred to act in meiotic recombination (Proudfoot & McCulloch, 2005b). The four other genes are designated *RAD51* paralogues; for more details, see Discussion, Section 7.2.2. To examine the role of two of these genes, *RAD51-3* and *RAD51-5*, gene disruption techniques were used to create homozygous mutants in *T. brucei*, and the resulting phenotypes were analysed (Proudfoot & McCulloch, 2005a). It was shown that the two *RAD51* paralogues, *RAD51-3* and *RAD51-5*, have a role in DNA repair and homologous recombination and in the case of *RAD51-3*, VSG switching (Proudfoot & McCulloch, 2005a). As there is comprehensive published data on the phenotypes of *rad51* *-/-*, *rad51-3* *-/-*, and *rad51-5* *-/-* mutants (McCulloch & Barry, 1999; Proudfoot & McCulloch, 2005a), for completeness, it was necessary to generate *rad51-4* *-/-* and *rad51-6* *-/-* mutants. In this chapter, the *RAD51-4* and *RAD51-6* mutant cell lines are generated and analyses of the resulting phenotypes are described. This will help to understand the role of the four *RAD51* paralogues and ascertain whether or not each act in DNA repair and recombination, as predicted by sequence homology.

To investigate the role of *RAD51-4* and *RAD51-6*, heterozygous and homozygous mutants were generated by gene disruption of the *RAD51-4* and *RAD51-6* ORFs, respectively. As *T. brucei* is diploid two rounds of transformations were required. To make comparable gene disruption mutants to those generated previously for *rad51* *-/-*, *rad51-3* *-/-*, and *rad51-5* *-/-*, the middle of each ORF was replaced by a different resistance marker. To do this, antibiotic resistance cassettes for the blasticidin S deaminase (*BSD*) gene of

*Streptomyces griseochromogenes*, encoding blasticidin resistance and the puromycin *N*-acetyl transferase (*PUR*) gene of *Streptomyces alboniger*, encoding puromycin resistance were integrated into the two alleles of each gene, generating first heterozygous mutants and then homozygous mutants. As a result of the integration, the Walker A and B boxes, conserved in all the *RAD51* paralogue proteins, were deleted. These mutants were generated in *T. brucei* bloodstream stage strain 3174.2 cells, to allow direct comparison with previous studies with the other *RAD51* paralogue mutants, *RAD51-3* and *RAD51-5*, which has been shown in Chapter 4 of this thesis. In particular this strain enables analysis of the VSG switching frequencies and mechanisms during antigenic variation (Hartley & McCulloch, 2008; McCulloch *et al.*, 1997; McCulloch & Barry, 1999; Rudenko *et al.*, 1996).

## 3.2 Generation of the *RAD51-4* gene disruption mutant

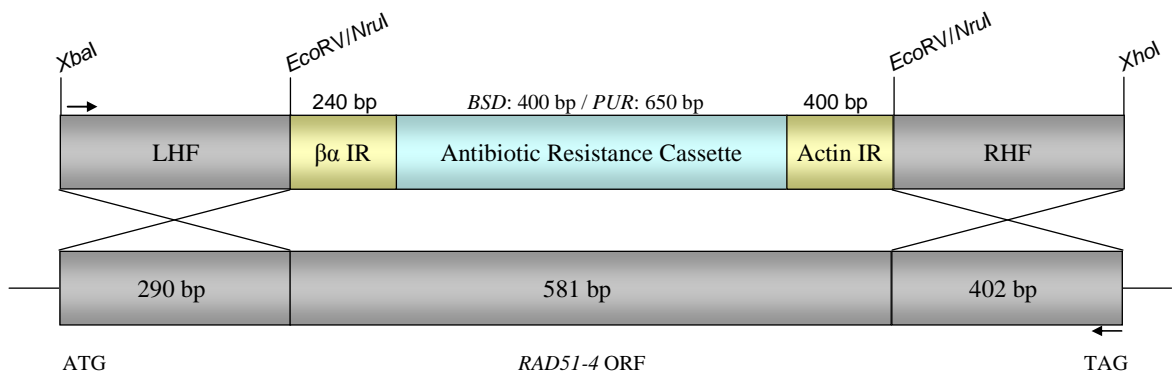
### 3.2.1 Design of the *RAD51-4* knockout constructs

The constructs for generating *RAD51-4* mutants were designed to allow homologous recombination of targeted homologous regions using the left hand flank (LHF) and the right hand flank (RHF) derived from the start and end of the *RAD51-4* ORF respectively. 581 bp of the centre of the *RAD51-4* ORF is replaced with antibiotic cassettes in this approach; see Figure 3.1. The nucleotide sequence of the *RAD51-4* ORF and the primer binding sites are shown in Appendix 1. The LHF of *RAD51-4* was PCR-amplified using primers *RAD51-4* StartORF For and *RAD51-4* StartORF Rev. The RHF of *RAD51-4* was PCR-amplified using primers *RAD51-4* EndORF For and *RAD51-4* EndORF Rev, (see Materials and Methods, Section 2.4, Table 2-1). The LHF and RHF PCR fragments, of 290 bp and 402 bp respectively were PCR-amplified from genomic DNA extracted from strain 3174.2 using Herculase (Stratagene) DNA polymerase enzyme, and were digested with *Xba*I and *Eco*RV, and *Xho*I and *Eco*RV respectively. The two fragments were cloned into *Xba*I and *Xho*I digested pBluescript KS (Stratagene) in a three way ligation reaction.

Antibiotic resistance cassettes containing the *BSD* (400 bp) and *PUR* (650 bp) ORFs were PCR-amplified from the plasmids pCP101 and pCP121 (Proudfoot & McCulloch, 2005a). Each resistance cassette is flanked by RNA processing sites: upstream  $\beta\alpha$ -tubulin (240 bp) and downstream actin (400 bp) intergenic regions. For each PCR-amplification, the primers Tubulin For and Actin Rev were used, (see Materials and Methods, Section 2.4, Table 2-1). These primers have blunt restriction enzyme sites, *Eco*RV for the *BSD* and

*NruI* for the *PUR* resistance cassettes. The resulting PCR fragments (approximately 1040 bp for the *BSD* cassette and 1340 bp for the *PUR* cassette) were cloned into Topo2.1 vector (pCR<sup>®</sup>2.1) (Invitrogen, see Materials and Methods, Section 2.4.7) and sequenced.

From the Topo clones, the antibiotic resistance cassettes were excised using blunt restriction enzymes, *EcoRV* for the *BSD* cassette and *NruI* for the *PUR* cassette. The excised constructs were then gel purified and ligated between the LHF and RHF of *RAD51-4* in pBluescript which had been digested with *EcoRV*. Insertion and correct orientation of the resistance cassettes between the LHF and RHF were confirmed by restriction mapping; the resulting constructs were sequenced and called  $\Delta$ *RAD51-4::BSD* and  $\Delta$ *RAD51-4::PUR*.



**Figure 3-1: *RAD51-4* gene disruption construct and gene disruption strategy.**

LHF: Left hand targeting flank has homology to the 5' region of the *RAD51-4* ORF. RHF: Right hand targeting flank has homology to the 3' region of the *RAD51-4* ORF. The 5' targeting flank (LHF) was 290 bp and the 3' targeting flank (RHF) was 402 bp. The antibiotic resistance cassettes (blasticidin: *BSD*/puromycin: *PUR*) have  $\beta\alpha$  tubulin intergenic region ( $\beta\alpha$  IR) and actin intergenic region (Actin IR) processing signals. The total length of the ORF for *RAD51-4* is 1273 bp and the sizes of individual the components (in bp) and restriction sites used in the cloning are indicated.

### 3.2.2 Generating the *RAD51-4* mutant cell lines

To generate *rad51-4*  $-/-$  mutants, two consecutive transformations were performed for each of the above constructs  $\Delta$ *RAD51-4::BSD* and  $\Delta$ *RAD51-4::PUR*. Two independent transformations using the construct  $\Delta$ *RAD51-4::BSD*, generated two independent *RAD51-4* heterozygous mutant cell lines respectively, which were named X $+/-$  and Y $+/-$ . Integration of the constructs into the expected gene locus was screened by PCR using a forward primer upstream of the ORF (Upstream of *RAD51-4* ORF For) and the reverse primer of the RHF (*RAD51-4* EndORF Rev) (data not shown). Following selection of the two clones, each was subjected to a second round of transformation using construct  $\Delta$ *RAD51-4::PUR* to generate homozygous mutant cell lines, X $-/-$  and Y $-/-$ . For each transformation, both constructs were linearised using the restriction enzymes *XbaI* and *XhoI* and then

phenol/chloroform extracted and ethanol precipitated.  $5 \times 10^7$  bloodstream stage *T. brucei* 3174.2 cells were transformed by electroporation in 0.2 cm cuvettes with approximately 5  $\mu\text{g}$  of each of the digested constructs (see Materials & Methods, Section 2.1.2). The electroporated cells were incubated for 24 hours in non-selective HMI-9 media to allow recovery. A volume of  $1 \times 10^7$  cells were diluted in 36 ml of selective HMI-9 media. The diluted cells were then plated over a 24 well plate at 1.5 ml per well, i.e.  $4.16 \times 10^5$  cells per well. Blasticidin (Calbiochem) at a concentration of  $2.5 \mu\text{g} \cdot \text{ml}^{-1}$  was used for selection of the heterozygous mutant cells following the first transformation and puromycin (Calbiochem) at a concentration of  $0.5 \mu\text{g} \cdot \text{ml}^{-1}$  was used to generate homozygous mutant cells following the second transformation.

The *RAD51-4* +/- and *rad51-4* -/- cell lines were confirmed by PCR, Southern, reverse transcriptase-PCR and western analysis (see Section 3.2.3, below). Each cell line was given a WCMP number and stabilates were made (see Materials and Methods, Section 2.1.2). For the following section, the heterozygous mutant cell lines are defined as +/- and the homozygous mutant cell lines are defined as -/-.

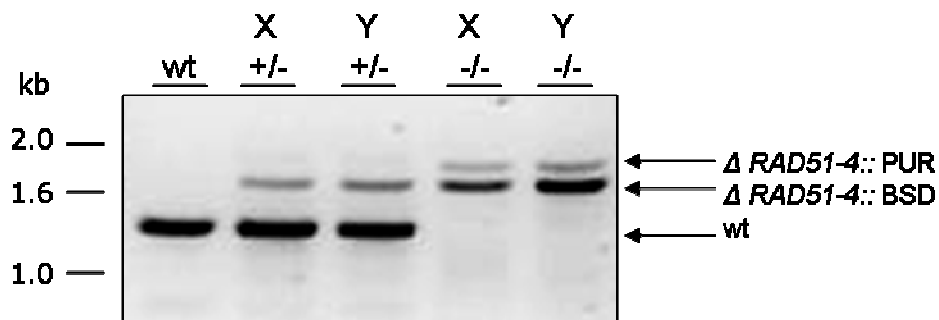
### 3.2.3 Confirmation of *RAD51-4* gene disruption mutants

#### 3.2.3.1 Confirmation of *RAD51-4* gene disruption by PCR

To determine if the *RAD51-4* locus had been disrupted PCR-amplification of the entire ORF, using a forward primer that bind upstream of the LHF (Upstream of *RAD51-4* ORF For, see Material and Methods, Section 2.4.1, Table 2-1) and a reverse primer that binds the 3' end of the RHF (*RAD51-4* EndORF Rev), was performed (see the black arrows in Figure 3-1). PCR-amplification was carried out using Taq DNA polymerase (Biolabs) and the reaction products were separated on a 1.0 % agarose gel. Figure 3-2 shows the PCR fragments obtained from the PCR reactions using genomic DNA extracted from the wildtype (wt), +/- and -/- cells. From the wt genomic DNA, one product of the size expected for the intact ORF was PCR-amplified (1273 bp). From the *RAD51-4* +/-, two products were PCR-amplified, one corresponding to the wildtype allele and the other to *BSD*-disrupted allele following integration into the ORF: integration of the  $\beta\alpha$ -tubulin-*BSD*-actin construct (1040 bp) creates an allele that is approximately 460 bp larger than the wt ORF. This size difference was observed. From the *rad51-4* -/-, no PCR product of the size of the wt allele was amplified. Instead, two larger PCR-products of sizes consistent with integration of both antibiotic cassettes were seen (*BSD* and *PUR*), disrupting both alleles: the  $\beta\alpha$ -tubulin-*PUR*-actin (1340 bp) construct is approximately 760 bp larger than



the wt ORF and 300 bp larger than the *RAD51-4::BSD* allele. These data indicate that the *RAD51-4* locus has been mutated as expected but cannot say if the *RAD51-4* gene has been copied to another part of the genome during transformation.



**Figure 3-2: Analysis of the *RAD51-4* gene disruption by PCR-amplification.**

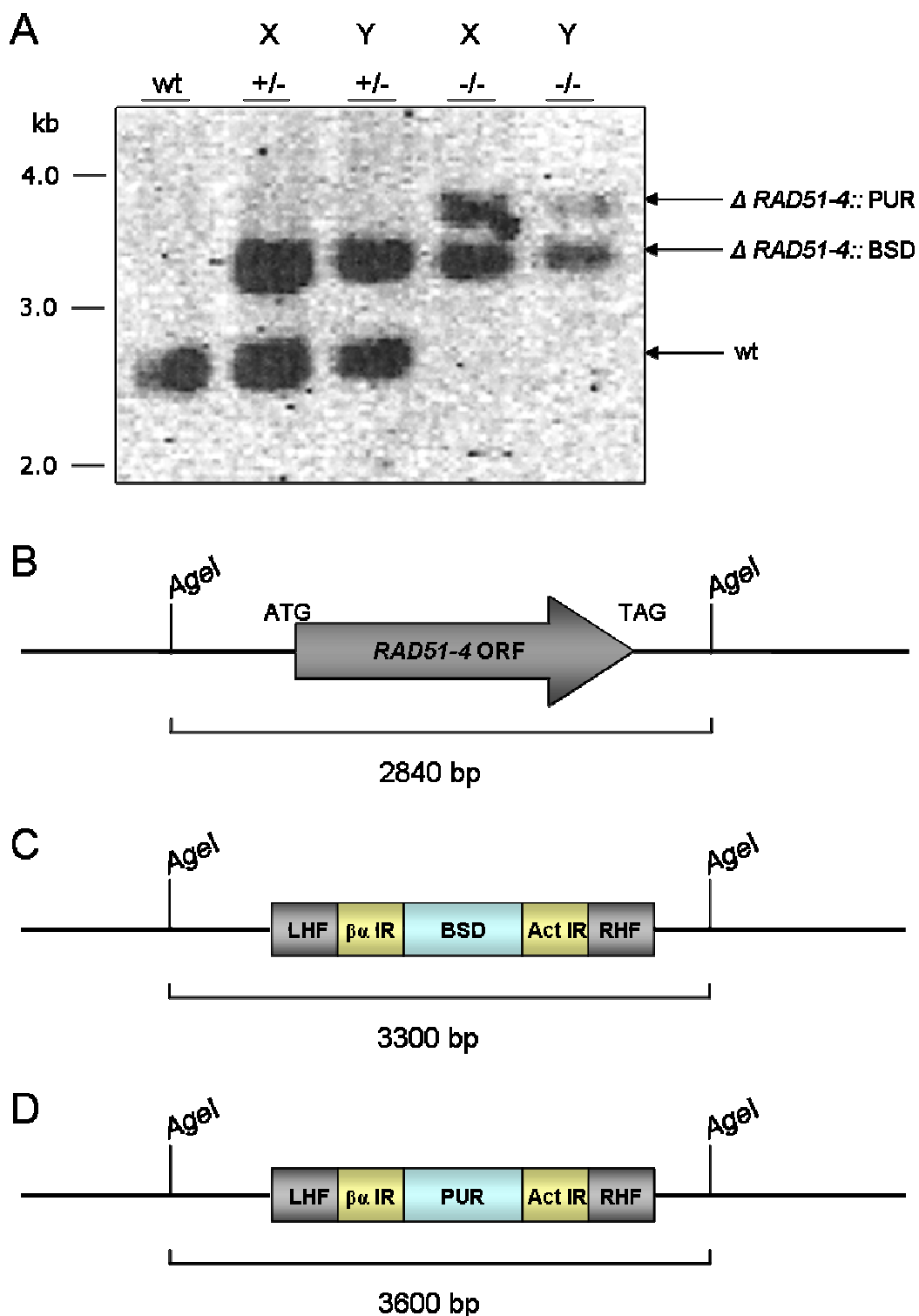
An agarose gel separation of the PCR products generated by PCR amplification of the *RAD51-4* locus using the forward primer: Upstream of *RAD51-4* ORF For and reverse primer: *RAD51-4* EndORF Rev is shown. PCR reactions were performed on genomic DNA from wildtype (wt) 3174.2 *T. brucei* cells, for two independent putative *RAD51-4* heterozygous mutants (X<sup>+/-</sup> and Y<sup>+/-</sup>) and two putative *rad51-4* homozygous mutants (X<sup>-/-</sup> and Y<sup>-/-</sup>). The DNA fragments of the sizes expected from the wt genes or from alleles disrupted by integration of the *BSD* resistance cassette ( $\Delta$ *RAD51-4::BSD*) or the *PUR* resistance cassette ( $\Delta$ *RAD51-4::PUR*) are indicated, as are the size markers.

### 3.2.3.2 Confirmation of *RAD51-4* gene disruption mutants by Southern blot analysis

Southern blot analysis was next carried out on the wildtype, *RAD51-4* <sup>+/</sup>- and *rad51-4* <sup>-/-</sup> cell lines, to confirm disruption of the ORF predicted by PCR amplification. Genomic DNA was extracted from each of the cell lines and digested with *Age*I, (see Material and Methods, Section 2.2.1), separated on a 0.8 % agarose gel and transferred on to a nylon membrane. The membrane was probed with the LHF of the ORF, part of the *RAD51-4* gene disruption constructs, *RAD51-4::BSD* and *RAD51-4::PUR*. The probe was made by PCR-amplifying the LHF using primers *RAD51-4* StartORF For and *RAD51-4* StartORF Rev. This resulting product was then gel extracted and labelled using the *Gene Images* CDP-*Star* Detection Module from Amersham, (see Materials and Methods, Section 2.5). The results of the Southern blot analysis are shown in Figure 3-3.

The Southern blot analysis shows that in wildtype cells, one fragment of the size expected (2840 bp) was obtained when the genomic DNA was digested with *Age*I. In the putative *RAD51-4* <sup>+/</sup>- mutants two products were obtained, one corresponded to the wildtype allele and the other of the size expected for the *BSD*-disrupted allele (3300 bp). In the *rad51-4* <sup>-/-</sup> mutants, no wildtype band fragment was observed but there were two larger products of the sizes corresponding to the *BSD*-disrupted (3300 bp) and *PUR*-disrupted alleles (3600 bp) (see Figure 3-3). As explained in Section 3.2.3.1 above, the  $\Delta$ *RAD51-4::BSD* construct

adds 460 bp to the wt *RAD51-4* allele, while the  $\Delta$ *RAD51-4::PUR* construct added 760 bp. These data confirm the PCR analysis of the *RAD51-4* locus, showing the ORF has been disrupted as expected. There was no apparent sign of the intact *RAD51-4* sequence in the *rad51-4*  $-/-$  cells, indicating these are mutated and no compensatory rearrangements have retained the *RAD51-4* gene.

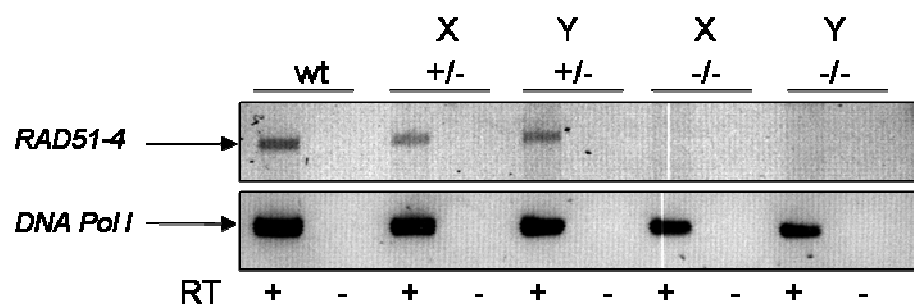


**Figure 3-3: Analysis of *RAD51-4* gene disruption by Southern blot.**

**Figure 3-3A:** Southern blot showing *AgeI*-digested genomic DNA and probed with the LHF of the *RAD51-4* ORF, see Figure 3-1. The extracted genomic DNA from wildtype (wt) 3174.2 *T. brucei* cells, from the two independent putative heterozygous mutants (X<sup>+/-</sup> and Y<sup>+/-</sup>) and from the two independent putative homozygous mutants (X<sup>-/-</sup> and Y<sup>-/-</sup>) was digested with *AgeI*, and separated on a 0.8 % agarose gel. The DNA was then transferred to a nylon membrane and probed with the LHF of the *RAD51-4* ORF. The DNA fragments of the size expected from the wt genes or from alleles disrupted by integration of the *BSD* resistance cassette (Δ*RAD51-4*::*BSD*) or the *PUR* resistance cassette (Δ*RAD51-4*::*PUR*) are indicated, as are the size markers. **Figure 3-3B** shows the locations of *AgeI* restriction enzyme sites around the *RAD51-4* ORF, predicted by sequence analysis (GeneDB).

### 3.2.3.3 Confirmation of *RAD51-4* gene disruption mutants by RT-PCR

To see whether or not *RAD51-4* mRNA was transcribed in each of the cell lines RT-PCR was carried out on cDNA generated from total RNA (see Material and Methods, Section 2.3). PCR-amplification of the entire ORF was attempted using a forward primer that binds to the LHF (*RAD51-4* StartORF For) and a reverse primer that binds to the RHF (*RAD51-4* EndORF Rev). Figure 3-4 shows the resulting PCR products separated on a 1.0 % agarose gel. In wildtype and *RAD51-4* +/- cell lines, one product of the size expected (1273 bp) was PCR-amplified from the cDNA. In the *rad51-4* -/- mutants, no *RAD51-4* mRNA was detected. A fragment of DNA polymerase I was amplified from the cDNA of all the cell lines, as confirmation that the RNA present from the extract was converted to cDNA. Control RT reactions performed without RT enzyme (RT-) confirmed that the PCR products amplified were not from contaminating genomic DNA which was not destroyed by the DNase treatment. This analysis confirmed that in the homozygous mutant cell lines *RAD51-4* mRNA is not detectably present.



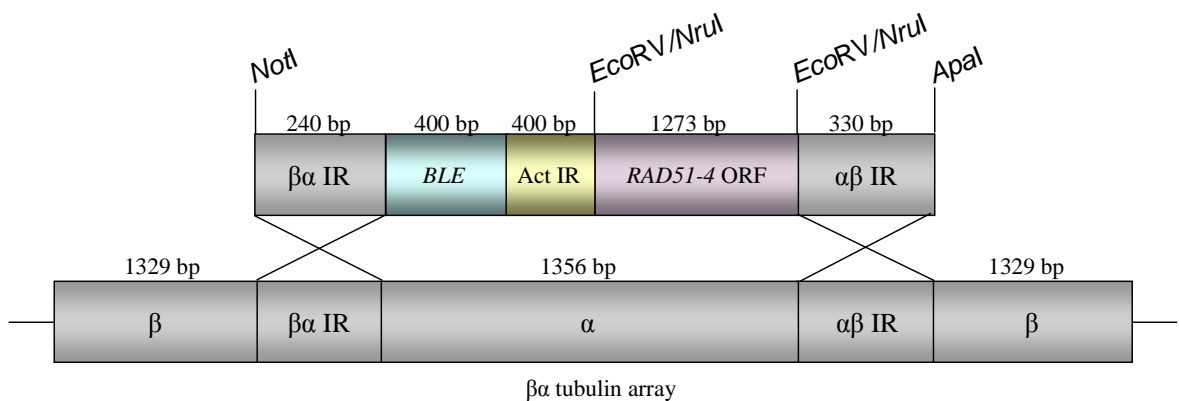
**Figure 3-4: Analysis of the expression of the *RAD51-4* mRNA in the mutant cell lines by RT-PCR amplification.**

Using the reverse transcriptase (RT) enzyme, cDNA was generated from mRNA extracted from the wildtype (wt) 3174.2 *T. brucei* cells, from the two independent putative *RAD51-4* heterozygous mutants (X +/- and Y +/-) and from the two independent putative *rad51-4* homozygous mutants (X -/- and Y -/-). This cDNA was used as template in the PCR reaction. PCR amplification was attempted using the forward primer (*RAD51-4* StartORF For) and reverse primer (*RAD51-4* EndORF Rev). The PCR products were then separated on a 1.0 % gel. As a positive control the DNA polymerase I ORF (DNA PolI) was amplified from the cDNA of all the mutant cell lines. RT + and - indicates the presence of RT enzyme (+) and the absence of RT enzyme (-).

## 3.3 Generation of *RAD51-4* re-expressor cell lines

Having confirmed that the *rad51-4* -/- mutants could be generated and are viable, it was necessary to re-express the *RAD51-4* in the mutant background, to show that the mutant phenotypes were due to the *RAD51-4* gene knockout. To this end, the entire *RAD51-4* ORF was PCR-amplified using the *RAD51-4* REX primers, (see Materials and Methods, Section 2.4.1, Table 2-1) and cloned using blunt restriction sites, *NruI*, into *EcoRV* digested pRM481 vector (McCulloch, gift). The pRM481 contains a phleomycin resistance cassette

(BLE; from *Streptoalloteichus hindustanus*) flanked upstream with a  $\beta\alpha$  tubulin intergenic region, and downstream with an actin intergenic region followed by a  $\alpha\beta$  tubulin intergenic region ( $\beta\alpha$ -BLE-Act- $\alpha\beta$ ). Following the cloning, the *RAD51-4* ORF is placed downstream of the actin intergenic region and upstream of the  $\alpha\beta$  tubulin intergenic sequences ( $\beta\alpha$ -BLE-Act-*RAD51-4*- $\alpha\beta$ ). This construct can be integrated into the  $\beta\alpha$  tubulin locus of *T. brucei*, and selection is provided by phleomycin. Recombination occurs on the  $\beta\alpha$  and  $\alpha\beta$  targeting regions, integrating the construct into the  $\beta\alpha$  tubulin array replacing the  $\alpha$  tubulin ORF, and allows expression of the *RAD51-4* mRNA through the flanking non-endogenous intergenic sequence (see Figure 3-5). This  $\beta\alpha$ -BLE-Act-*RAD51-4*- $\alpha\beta$  construct was digested with *NotI* and *ApaI* and transformed into the *rad51-4*  $-/-$  mutant cell lines and were allowed to recover for 24 hours in non-selective media. These were then clonally plated on  $2.5 \mu\text{g.ml}^{-1}$  of phleomycin (Calbiochem).



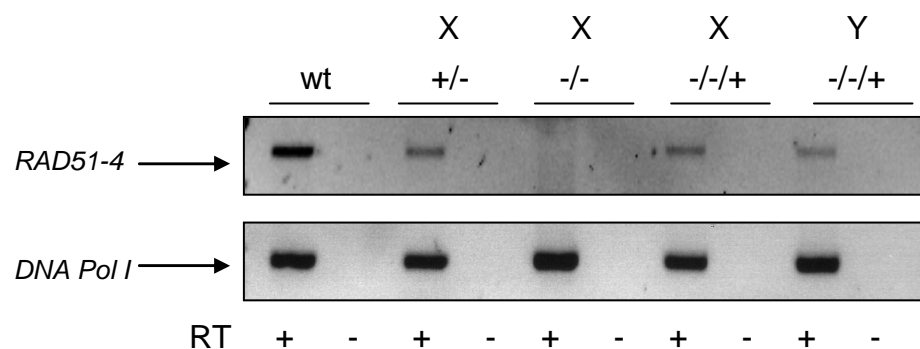
**Figure 3-5: Strategy for re-expression of the *RAD51-4* by recombination into the  $\beta\alpha$  tubulin array.**

The ORF of *RAD51-4* was cloned into the restriction site *NruI* of the pRM481 (top part of the diagram). *NotI* and *ApaI* digestion liberates the construct which was then transformed into the *rad51-4*  $-/-$  cell line. Recombination occurs on the  $\beta\alpha$  and  $\alpha\beta$  targeting regions, integrating the construct including the *RAD51-4* ORF into the  $\beta\alpha$  tubulin array replacing one of the  $\alpha$  tubulin ORFs.

### 3.3.1 Confirmation of the *RAD51-4* re-expressor cell lines by PCR and RT-PCR

Possible *RAD51-4* re-expressor clones following transformation of *rad51-4*  $-/-$  cell lines with the pRM481 re-expressor construct ( $\beta\alpha$ -BLE-Act-*RAD51-4*- $\alpha\beta$ ) were screened using PCR primers that bind to the LHF and RHF of *RAD51-4* (*RAD51-4* StartORF For and *RAD51-4* EndORF Rev, see Materials and Methods, Section 2.4.1). These primers PCR-amplified the integrated intact ORF and also the disrupted ORFs meaning that the re-expressor cell lines had three amplified fragments: one represented the re-expressor *RAD51-4* ORF, and the other two bands corresponded to the *BSD* and *PUR* disrupted alleles. From this, putative re-expressors were chosen for each independently generated

*rad51-4*  $-/-$  mutant (data is not shown), and were analysed in more detail. Firstly, RT-PCR-amplification analysis, using the primers that bind to the *RAD51-4* LHF and RHF (*RAD51-4* StartORF For and *RAD51-4* EndORF Rev) were used to detect the intact mRNA in the putative re-expressor ( $-/-/+$ ), see Materials and Methods, Section 2.3). In wildtype (wt), *RAD51-4*  $+/-$  (X $+/-$ ) and *RAD51-4* re-expressor (X $-/-/+$  and Y $-/-/+$ ) cell lines, one product of the size expected was amplified from the cDNA. In the *rad51-4*  $-/-$  (X $-/-$ ), no product was amplified from the cDNA. DNA polymerase I was amplified from the cDNA of all the cell lines, confirming that extracted RNA was converted to cDNA. The RT negative control has no RT enzyme (-) in the reaction and confirms that the PCR products generated arose from the cDNA and not from contaminating gDNA (Figure 3-6).



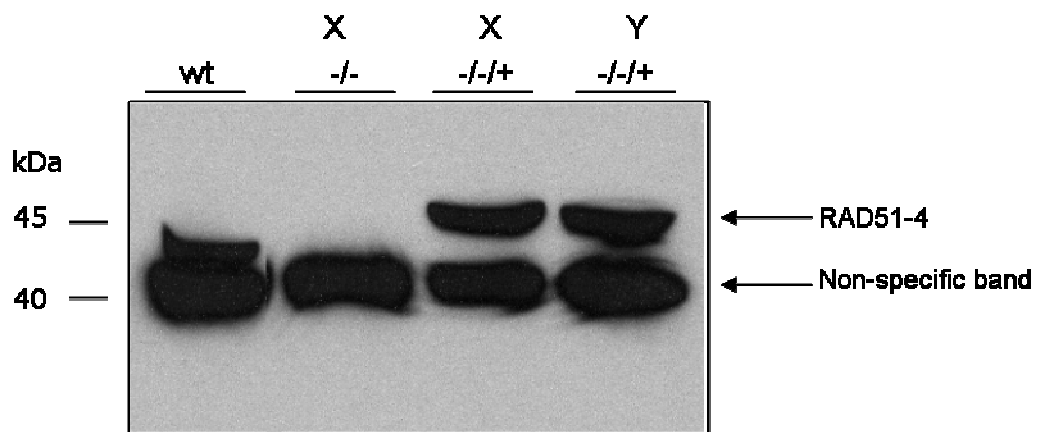
**Figure 3-6: Analysis of the expression of the *RAD51-4* mRNA in the *RAD51-4* re-expressor cell lines by RT-PCR amplification.**

Using the reverse transcriptase (RT) enzyme, cDNA was generated from mRNA extracted from the wildtype (wt) 3174.2 *T. brucei* cells, from the one putative *RAD51-4* heterozygous mutants (X $+/-$ ), from one putative *rad51-4* homozygous mutants (X $-/-$ ) and from the two independently generated *RAD51-4* re-expressor cell lines. This cDNA was used as template in the PCR reaction. PCR amplification was attempted using the forward primer (*RAD51-4* StartORF For) and reverse primer (*RAD51-4* EndORF Rev). The PCR products were then separated on a 1.0 % gel. As a positive control the DNA polymerase I ORF (DNA PolI) was amplified from the cDNA of all the mutant cell lines. RT + and - indicates the presence of RT enzyme (+) and the absence of RT enzyme (-).

### 3.3.2 Confirmation of the *RAD51-4* re-expressors by Western blot analysis

To determine if the *RAD51-4* protein was expressed in the *RAD51-4* re-expressor cell lines ( $-/-/+$ ), western blot analysis was carried out, using purified anti-*RAD51-4* anti-serum, (see Section 2.6.2.1). Whole cell extracts from wt, *rad51-4*  $-/-$  (X $-/-$ ) and *RAD51-4* re-expressors (X $-/-/+$  and Y $-/-/+$ ) were separated on a 10 % SDS-PAGE gel and transferred to a nylon membrane (see Materials and Methods, Section 2.6.3). The blot was probed with rabbit anti-*RAD51-4* anti-serum primary antibody at a dilution of 1:1000, and was detected with HRP-conjugated anti-rabbit IgG antibody (Molecular Probes) at a dilution of 1:5000. This data is shown in Figure 3-7. In wildtype cells (wt) two bands were seen, one of the

size expected for the RAD51-4 (45 kDa) and another smaller band of ~40 kDa which was non-specific binding of the anti-serum. In the homozygous mutant analysed (X<sup>-/-</sup>), no wildtype band of 45 kDa indicating that RAD51-4 is absent. The protein levels look comparable in the cell lines as the non-specific band appears to be similar in all the lanes. In both the putative re-expressor cell lines, the two bands were seen, one of these was the non-specific band, while the other was ~45 kDa and is therefore RAD51-4 expressed from tubulin. However it appears that this is slightly bigger than the species in the wildtype cells. It is unknown why the re-expressed RAD51-4 is larger than the wildtype RAD51-4. This may be due to the location of the ORFs. The endogenous *RAD51-4* may differ in *trans*-splicing compared to the transcribed *RAD51-4* gene from the  $\beta\alpha$  tubulin array (see Section 1.1.5). It could also be the case that the slightly bigger re-expressed RAD51-4 may have been modified post-translation, for example, by phosphorylation.



**Figure 3-7: Analysis of the RAD51-4 protein levels in the *RAD51-4* re-expressor cell lines by western blot analysis.**

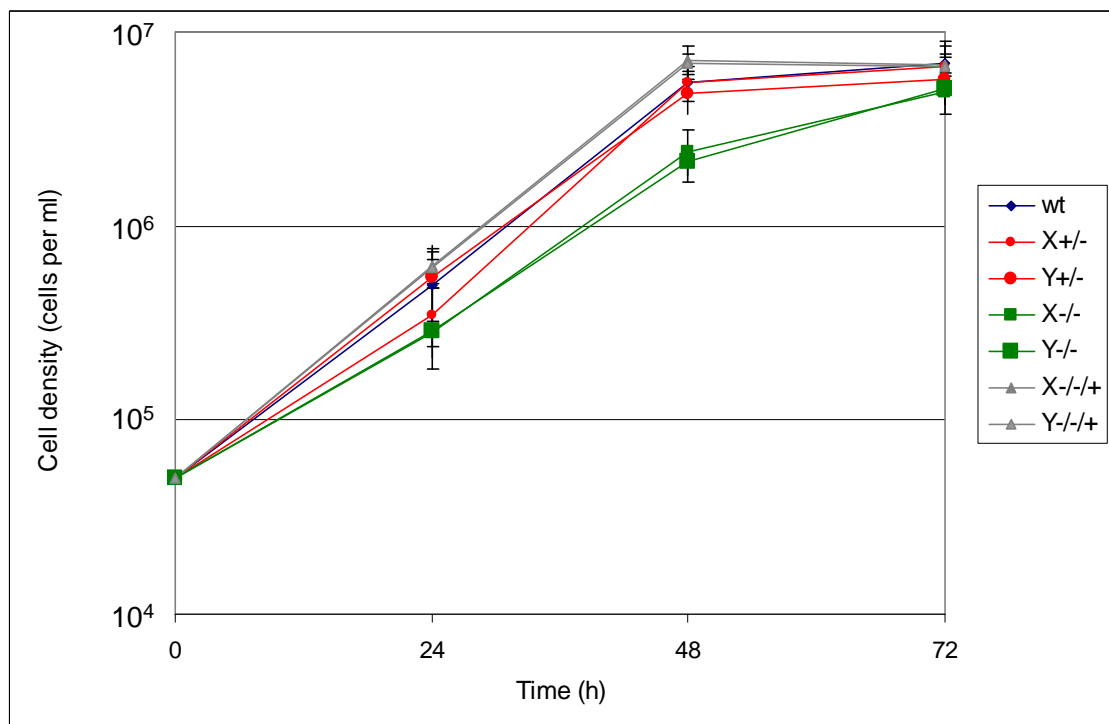
Western blot analysis showing the cell line extracts probed with anti-RAD51-4 anti-serum. Protein extracts from the wildtype (wt), *rad51-4* <sup>-/-</sup> and the re-expressor cell lines (X<sup>-/-</sup> and Y<sup>-/-</sup>) were separated on a 10 % SDS-PAGE gel, and then transferred to a nylon membrane. This was then probed with rabbit anti-RAD51-4 anti-serum primary antibody at a dilution of 1:1000, and was detected with HRP-conjugated anti-rabbit IgG antibody at a dilution of 1:5000. The predicted sizes are indicated in kDa.

### 3.4 Phenotypic analysis of *RAD51-4* mutants and re-expressor cell lines

Once the *RAD51-4* mutant cell lines were generated and confirmed, phenotypic analysis of the cell lines was carried out. The phenotypes observed for the homozygous mutant cell lines were compared to that of the wildtype and heterozygous mutant cell lines. To this end, growth rate, DNA-damage sensitivities, ability to form RAD51 foci and VSG switching frequency were analysed. The following sections described these results.

### 3.4.1 Analysis of *in vitro* growth rate of *RAD51-4* mutant cell lines

To determine if knockout of *RAD51-4* has any gross effects on the growth of *T. brucei*, population doubling time was measured. Since mutation of the mammalian RAD51 paralogues, and *RAD51-3* and *RAD51-5* of *T. brucei* caused reduced growth rates, similar growth impairment may be expected (Proudfoot & McCulloch, 2005a; Takata *et al.*, 2001). To examine this, wildtype (wt), heterozygous (+/-) and homozygous (-/-) mutants were inoculated (in triplicate) at  $5 \times 10^4$  cells per ml in HMI-9 media and cell densities were counted at fixed time points subsequently (see Materials and Methods, Section 2.8.1). The results were plotted on a semi-logarithmic graph (see Figure 3-8), and the doubling time for both the *RAD51-4* +/- (X, Y) and *rad51-4* -/- (X, Y) mutants were calculated and compared to that of the wt (see Table 3-1). The same analysis was conducted with the *RAD51-4* -/-/+ (X, Y), where the *RAD51-4* ORF has been integrated into the homozygous mutants.



**Figure 3-8: Analysis of growth of *RAD51-4* mutants *in vitro*.**

A semi-logarithmic plot of cell densities against fixed time points (24 h, 48 h, and 72 h) is shown. Wildtype strain 3174.2 (wt) is compared with two independent *RAD51-4* heterozygous mutants (X+/-, Y+/-), two independent *rad51-4* homozygous mutants (X-/-, Y-/-) and *RAD51-4* re-expressor cell lines (X-/-/+, Y-/-/+). The graph shows the means of triplicate data with the 95 % confidence intervals indicated.



**Table 3-1: *in vitro* doubling times for *RAD51-4* mutant cell lines.**

The population doubling time for the two independently generated *RAD51-4* heterozygous (X+/-, Y+/-), *rad51-4* homozygous (X-/-, Y-/-) and the *RAD51-4* re-expressor cell lines (X-/-/+, Y-/-/+) are compared with the doubling time for wildtype strain 3174.2 (wt). This data are the means of triplicate experiments (in hours) with the 95 % confidence intervals indicated (in brackets).

wt	X +/-	Y +/-	X -/-	Y -/-	X -/-/+	Y -/-/+
7.23 (+/-0.24)	7.12 (+/-0.46)	7.43 (+/-0.47)	9.03 (+/-0.55)	9.08 (+/-0.37)	6.78 (+/-0.25)	6.82 (+/-0.17)

**Table 3-2 : Statistical analysis of the *in vitro* doubling times for *RAD51-4* mutants.**

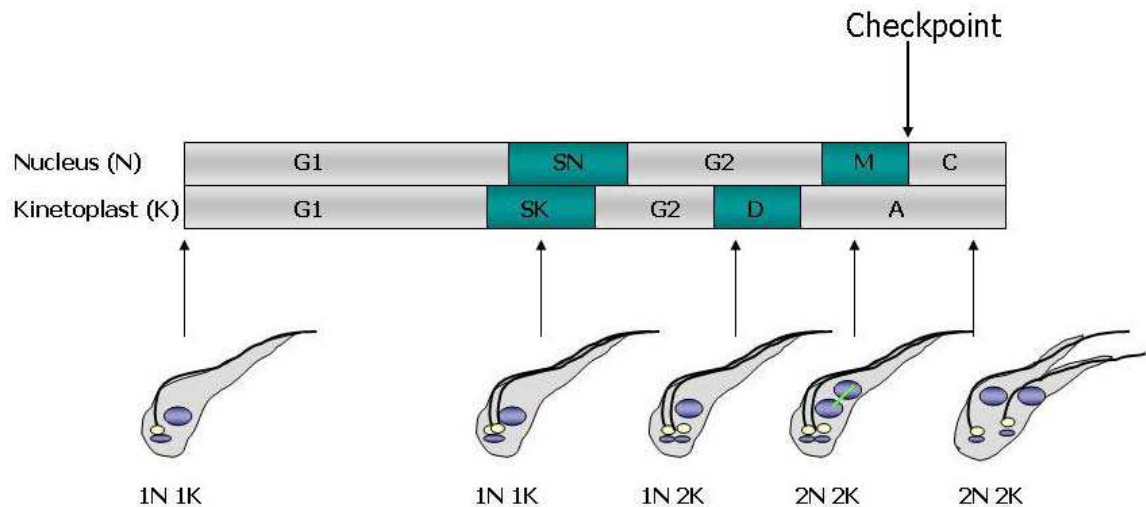
Unpaired, two-tailed Student's *t*-tests were carried out to compare the mean of the triplicate doubling time results for the *RAD51-4* heterozygous (X+/-, Y+/-), *rad51-4* homozygous (X-/-, Y-/-) and *RAD51-4* re-expressor cell lines (X-/-/+, Y-/-/+) against the wildtype (wt). Significant differences between the means is shown by a *p*-value of  $p \leq 0.05$ ; Not significantly different results are shown as "--". [NA is not applicable]

	wt	X +/-	Y +/-	X -/-	Y -/-	X -/-/+	Y -/-/+
wt	NA	-	-	0.0041	0.0012	-	-

Table 3-2 shows that the doubling times for each independent homozygous mutant is significantly different to that of the wildtype. This analysis was carried out using unpaired, two-tailed Student's *t*-test, and all other pair-wise comparisons of the other cell lines are shown in Appendix 3, Table 1.

These data show that *rad51-4* -/- mutants had a reduced growth rate, with each independent homozygous mutant, X-/- and Y-/-, having a decrease in growth of 44.3 % and 39.8 % compared to the wildtype at 48 h (see Figure 3-8). As a result, there was also an increased doubling time compared to the wildtype and *RAD51-4* heterozygous mutants: X-/- had a doubling time of 9.03 h (+/- 0.55) and Y-/- had a doubling time of 9.08 h (+/- 0.37) compared with the wildtype doubling time of 7.23 h (+/- 0.24); see Table 3-1. Re-addition of functional *RAD51-4* back into the homozygous mutant cell line enabled the cells to recover from the growth deficiency. This confirmed that lack of *RAD51-4* was responsible for the slow growth of the cell lines. The basis for this slow growth was not explored further, but could be due to increased cell death in the population or increased length of the cell cycle in each cell of the population. A growth phenotype due to increased cell death could be analysed by clonal survival assays similar to those described in Section 3.4.2.1 and Section 3.7.2.1. An increase in population doubling times, due to the cells taking longer to complete some stage(s) of the cell cycle or due to a cell cycle stall, could be analysed by quantification of the kinetoplast and nucleus. The nuclear and kinetoplast DNA replication and segregation are asynchronous, leading to differing copy numbers over the cell cycle (see Figure 3-9; McKean, 2003), which can be analysed microscopically by

DAPI staining. If there is a cell cycle defect in the mutant population, the percentage of cells in a specific cell cycle stage will differ from the wildtype.



**Figure 3-9: Replication and division of the nucleus and kinetoplast at specific cell cycle stages of *T. brucei* bloodstream form.**

*T. brucei* cells contain 1 nucleus (large purple circle) and 1 kinetoplast (small purple circle) (1N 1K) during G1 phase of the cell cycle. Synthesis of the kinetoplast DNA (SK) starts shortly before the nuclear DNA synthesis (SN). Kinetoplast division (D) occurs before nuclear division, resulting in the cells having 1 nucleus and 2 kinetoplasts (1N 2K). Nuclear division or mitosis (M) results in the cells having 2 nuclei and 2 kinetoplasts (2N 2K), and occurs before cytokinesis (C). The apportioning (A) phase of the cell cycle is when single copy organelles separate, for example basal bodies (white circles). Cytokinesis generates two daughter cells containing 1N 1K. Taken from C. Hartley Thesis (2008) and adapted from McKean (2003).

### 3.4.2 Analysis of DNA-damage sensitivity of *RAD51-4* mutant cell lines

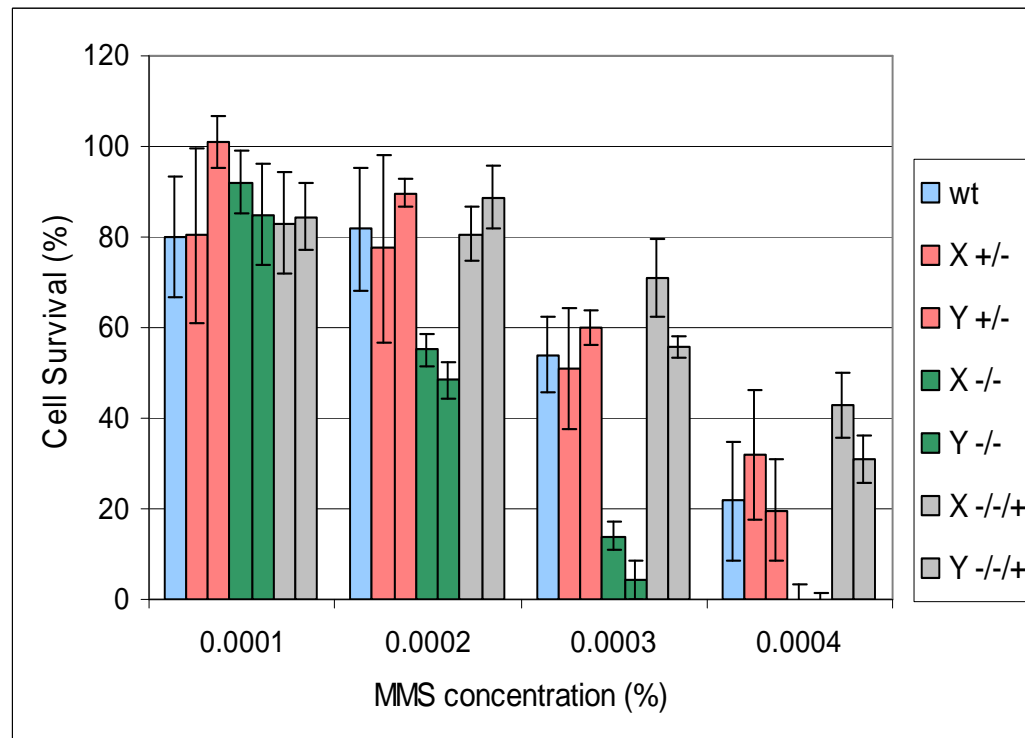
To analyse if disruption of the *RAD51-4* ORF leads to increased DNA-damage sensitivity, growth of the mutant cell lines was compared following treatment with two different DNA damaging agents, methyl methanesulfonate (MMS) and phleomycin. MMS (Sigma) is a DNA alkylating agent which modifies two DNA bases: guanine (to 7-methylguanine) and adenine (to 3-methyladenine) (Beranek, 1990; Lundin *et al.*, 2005). Historically, MMS was used as a 'radiomimetic' compound in yeast studies because many of the original radiation sensitive (rad) mutants were found to be sensitive to the compound (Wyatt & Pittman, 2006). It is still used to induce mutagenesis due to its DNA-damaging properties. It is also exploited in recombination experiments, as it is known to cause DNA-damage induced sister chromatid exchange (SCE) (Fasullo *et al.*, 1994). DNA-damage induced by MMS results in single strand and double strand breaks (Lundin *et al.*, 2005), although it is not known whether MMS directly causes DNA strand breaks, or whether the strand breaks observed result from events subsequent to the DNA damage for instance by triggering base excision repair or causing stalling/collapse of replication forks (Wyatt & Pittman, 2006).

Phleomycin (Calbiochem) is a glycopeptide and a member of the bleomycin group of antibiotics and anti-neoplastics (Dumas *et al.*, 1994). Bleomycin causes both single strand and double strand breaks, with the latter being the most predominant, due to the ability of the compound to intercalate into DNA (Claussen & Long, 1999). It acts through the oxidation of a DNA-Fe(II)-bleomycin complex, with the formation of free radicals (Giloni *et al.*, 1981). This oxidation of Fe(II) and reduction of oxygen when bound to pyrimidine nucleotides induces the DNA damage and can result in chromosomal aberrations (Claussen & Long, 1999). DNA cleavage by Fe(II)-bleomycin has also been shown to be a selective process of strand-scission, occurring preferentially at guanine-pyrimidine base sequence (5'-GC and 5'-GT) (Claussen & Long, 1999).

Two assays were used to ascertain whether or not the *rad51-4* *-/-* mutants display increased sensitivity to phleomycin or MMS induced DNA-damage. The first assay was clonal survival, measuring the outgrowth of the population from a single cell in the presence of the two different DNA-damaging agents. The second was assaying the sensitivity to the DNA-damaging agents by measuring absorbance of alamar blue and calculating IC<sub>50</sub> values for each cell line. IC<sub>50</sub> is a measure of the effectiveness of a compound, in this case DNA-damaging agents MMS and phleomycin, in inhibiting biological survival or growth. It is a quantitative measure and indicates how much of the DNA-damaging agents are required to inhibit metabolic capability by 50 %. The IC<sub>50</sub> values of the DNA-damaging agents were determined by constructing a dose-response curve (GraphPad Prism 4).

#### **3.4.2.1 Clonal survival of RAD51-4 mutant cell lines**

Clonal survival assay was performed by diluting mid-log phase ( $2 \times 10^6$  cells per ml) cell lines to a concentration of 1 cell per well over a 96 well plate, measuring the number of cells that grew out to an observable population with varying concentrations of MMS and phleomycin, (see Materials and Methods, Section 2.8.2.1 for more details). Each drug concentration was plated in triplicate and growth in the well was scored microscopically up to 14 days after cloning. This was graphed to show cell survival as a percentage of clonal growth when compared with non-treated cells. Figure 3-10 shows the clonal survival of *RAD51-4* mutant cell lines compared to that of the wildtype cells at increasing concentrations of MMS as follows: 0.0001 %, 0.0002 %, 0.0003 %, and 0.0004 %.



**Figure 3-10: Clonal survival of *RAD51-4* mutants *in vitro* when treated with methyl methanesulfonate (MMS).**

Cell survival is shown as percentage of clonal growth when treated with MMS compared to growth in the absence of MMS. Each cell line, wildtype strain 3174.2 (wt), two independent *RAD51-4* heterozygous mutants (X+/-, Y+/-), *rad51-4* homozygous mutants (X-/-, Y-/-) and *RAD51-4* re-expressor cell lines (X-/-/+, Y-/-/+) was treated with increasing concentration of MMS (0.0001 %, 0.0002 %, 0.0003 %, 0.0004 %). The graph shows the means of triplicate data, with 95 % confidence intervals indicated by the error bars.

With increasing concentrations of MMS, from 0.0001 % to 0.0004 %, there was clear evidence for decreased survival of each independent *rad51-4* -/- mutant, X-/- and Y-/-, compared to the other cell lines. At a MMS concentration of 0.0003 % this distinction was clearest: wildtype survival was 53.9 % (+/- 13.2) of untreated cells, whereas *rad51-4* X-/- mutant survival was 14 % (+/- 3.3) and *rad51-4* Y-/- was 4.3 % (+/- 1.4). Analysis of the results by the Student's *t*-tests for clonal survival at this MMS concentration (0.0003 %) is shown in Table 3-3. This confirms that clonal survival of each independent *rad51-4* -/- mutant is significantly different to that of the wildtype cell line.

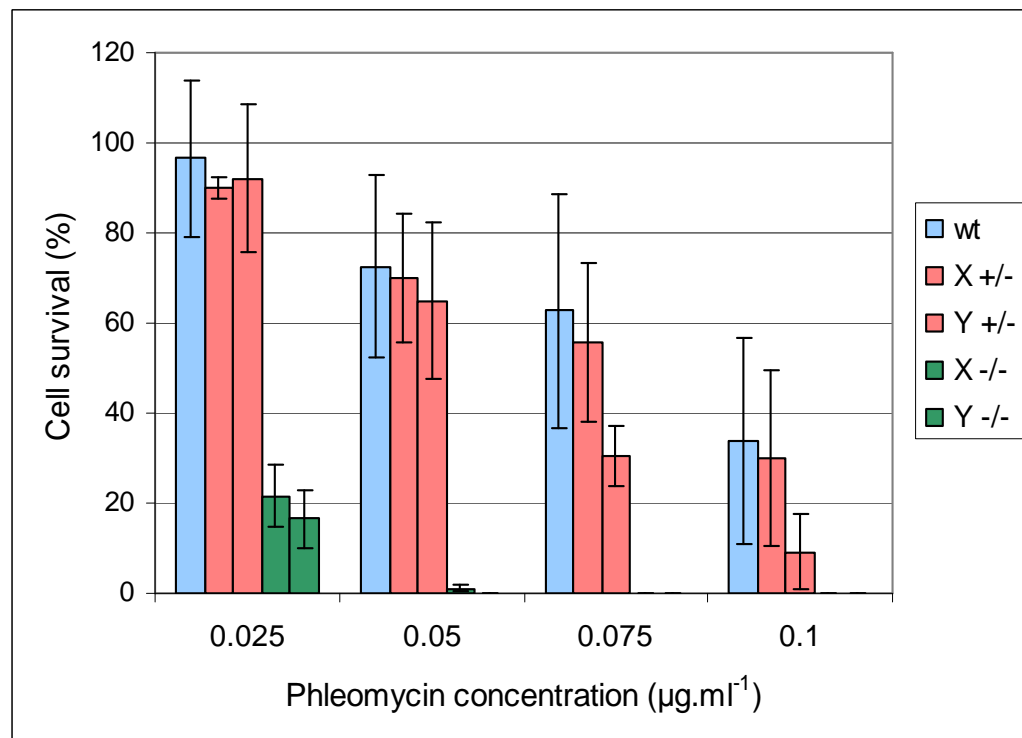
**Table 3-3: Statistical analysis of the clonal survival of *RAD51-4* mutants when treated with 0.0003 % methyl methanesulfonate (MMS).**

Unpaired two-tailed Student's *t*-tests were carried out to compare the mean of the triplicate clonal survival results for the cell lines against the wildtype. Significant difference between the results is shown by a *p*-value of  $p \leq 0.05$ , while not significantly different results are shown as "-". [NA is not applicable] Each cell line is indicated as follows: wildtype strain 3174.2 (wt), two independent *RAD51-4* heterozygous mutants (X+/-, Y+/-), *rad51-4* homozygous mutants (X-/-, Y-/-) and *RAD51-4* re-expressor cell lines (X-/-/+, Y-/-/+).

	wt	X +/-	Y +/-	X -/-	Y -/-	X -/-/+	Y -/-/+
wt	NA	-	-	0.0349	0.0420	-	-

Student's *t*-tests to compare the mean of the triplicate clonal survival results for the cell lines against each other at the other concentrations of MMS, 0.0001 %, 0.0002 % and 0.0004 %, were also carried out, and are shown in Appendix 3, Table 2. These results confirm that with increasing concentrations of MMS, there is a significant reduction of clonal survival of the *rad51-4* *-/-* cells compared with the wildtype and *RAD51-4* *+/-*. With the addition of a functional *RAD51-4* gene back into the homozygous mutants, the percentage of survival is increased, indicating this phenotype is a consequence of *RAD51-4* mutation.

Similar analysis was carried out using phleomycin as a DNA damaging agent. Figure 3-11 shows the clonal survival of *RAD51-4* mutant cell lines compared with that of the wildtype cells when treated with four increasing concentrations of the drug: 0.025  $\mu\text{g.ml}^{-1}$ , 0.05  $\mu\text{g.ml}^{-1}$ , 0.075  $\mu\text{g.ml}^{-1}$ , and 0.1  $\mu\text{g.ml}^{-1}$ .



**Figure 3-11: Clonal survival of *RAD51-4* mutants *in vitro* when treated with phleomycin.**

Cell survival is shown as percentage of clonal growth when treated with phleomycin compared to growth in the absence of drug. Each cell line wildtype 3174.2 (wt), two independent *RAD51-4* heterozygous mutants (X $+/$ -, Y $+/$ -), and *rad51-4* homozygous mutants (X $-/-$ , Y $-/-$ ) was treated with increasing concentration of phleomycin (0.025  $\mu\text{g.ml}^{-1}$ , 0.05  $\mu\text{g.ml}^{-1}$ , 0.075  $\mu\text{g.ml}^{-1}$ , 0.1  $\mu\text{g.ml}^{-1}$ ). The graph shows the means of triplicate data, with the 95 % confidence intervals indicated by error bars.

As was seen for MMS, increasing concentrations of phleomycin, from 0.025  $\mu\text{g.ml}^{-1}$  to 0.1  $\mu\text{g.ml}^{-1}$ , caused a clear decrease of survival of each independent *rad51-4* *-/-* X and Y, compared with the wt and *RAD51-4* *+/-*. At a phleomycin concentration of 0.025  $\mu\text{g.ml}^{-1}$ , wildtype survival was 96.5 % ( $\pm$  17.3) compared with 21.6 % ( $\pm$  6.8) and 16.8 % ( $\pm$  6.8).

6.4) for the *rad51-4*  $-/-$  mutants, X $-/-$  and Y $-/-$  respectively. The results of the Student's *t*-tests for clonal survival at a phleomycin concentration of  $0.025 \mu\text{g.ml}^{-1}$  are shown in Table 3-4, confirming that the clonal survival of each independent homozygous mutant was significantly reduced compared to that of the other cell lines.

**Table 3-4: Statistical analysis of the clonal survival of *RAD51-4* mutants when treated with  $0.025 \mu\text{g.ml}^{-1}$  of phleomycin.**

Unpaired, two-tailed Student's *t*-tests were carried out to compare the mean of the triplicate clonal survival results for the cell lines against the wildtype (wt), significant difference between the results is shown by a *p*-value of  $p \leq 0.05$ , whereas not significantly different results are shown as "-". [NA is not applicable] Each cell line is described: wildtype strain 3174.2 (wt), two independent *RAD51-4* heterozygous mutants (X $+/+$ , Y $+/+$ ), and *rad51-4* homozygous mutants (X $-/-$ , Y $-/-$ ).

	wt	X $+/+$	Y $+/+$	X $-/-$	Y $-/-$
wt	NA	-	-	0.0412	0.0306

Unpaired, two-tailed Student's *t*-tests to compare the means of the triplicate clonal survival results for the cell lines against each other at the different concentrations of phleomycin,  $0.05 \mu\text{g.ml}^{-1}$ ,  $0.075 \mu\text{g.ml}^{-1}$ , and  $0.1 \mu\text{g.ml}^{-1}$  are shown in Appendix 3, Table 3. At concentration of  $0.05 \mu\text{g.ml}^{-1}$  and  $0.075 \mu\text{g.ml}^{-1}$  of phleomycin, there was also a significant decrease of clonal survival of the *rad51-4*  $-/-$  mutants compared with the wt and *RAD51-4*  $+/+$  mutants. Note that the *RAD51-4* re-expressor cell lines were not analysed, as reintegration of *RAD51-4* was selected using phleomycin resistance.

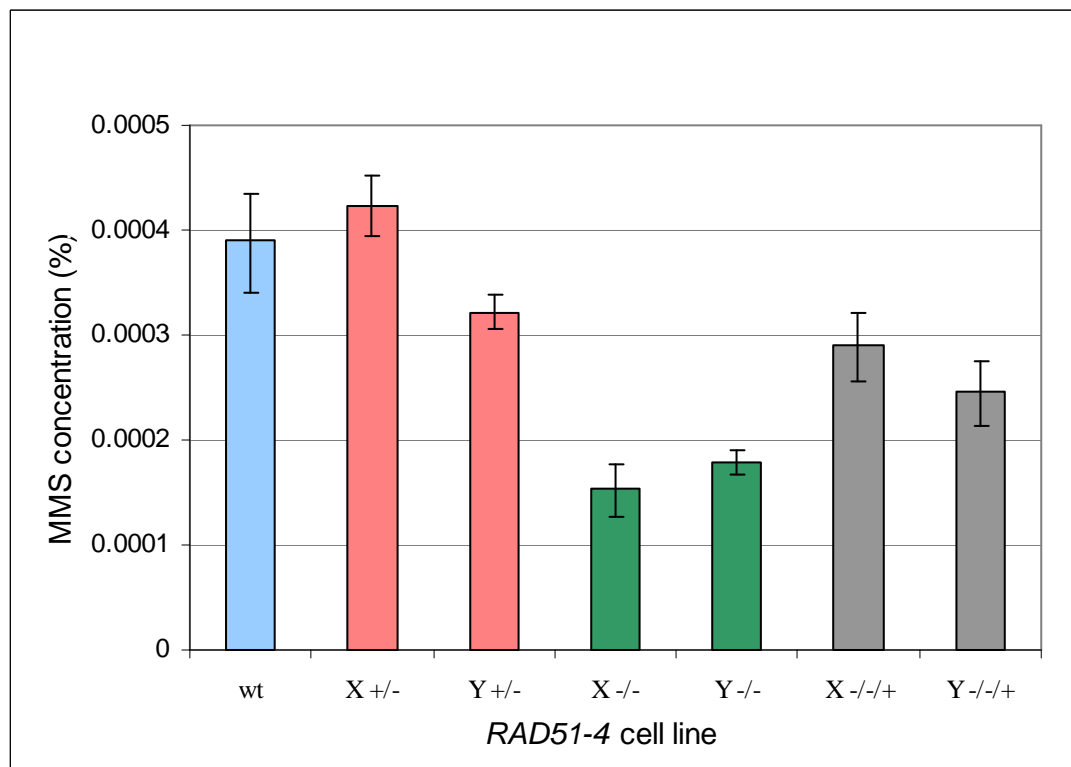
Taken together, these data show that the *rad51-4*  $-/-$  mutants are more sensitive to DNA-damaging agents, MMS and phleomycin, than the wildtype and *RAD51-4*  $+/+$  mutants. Most likely, this is due to impairment in the ability of the cells to repair these types of induced DNA damage.

### **3.4.2.2 *IC*<sub>50</sub> values of MMS and Phleomycin for *RAD51-4* mutant cell lines**

In order to quantify the extent of *rad51-4*  $-/-$  mutants DNA damage sensitivity, *IC*<sub>50</sub> values of the different cell lines were determined after MMS and phleomycin treatment. This was performed using an Alamar blue assay (see Materials and Methods, Section 2.8.2.2). Resazurin (or Alamar blue) is an indicator redox dye and is a quantitative measure of proliferating cells. The assay is based on the ability of viable, metabolically active cells to reduce resazurin to resorufin (Anoopkumar-Dukie *et al.*, 2005). This reduction results in a colour change from blue (resazurin) to pink (resorufin), which can be measured colorimetrically. This compound is non-toxic to cells and stable in the HMI-9 culture

medium, allowing measurement of cell proliferation *in vitro*. If the DNA-damaging agents impair the cell viability and proliferation, the capacity of cultures to reduce resazurin is also reduced. The rate of dye reduction is directly proportional to the number of viable cells present in the culture. Therefore, as a direct measure of the metabolic competence of cell cultures, colour change provides a convenient indicator of cell proliferation following DNA-damage which then allows the  $IC_{50}$  values to be calculated.

Figure 3-12 shows the calculated mean  $IC_{50}$  values for MMS of wildtype, *RAD51-4* +/-, *rad51-4* -/- and *RAD51-4* re-expressor cell lines. The wildtype cells' mean  $IC_{50}$  value was 0.000389 % (+/- 0.000048) which was approximately 2.5 fold higher compared than the two independent *rad51-4* -/- mutants: 0.000154 % (+/- 0.000025) and 0.000178 % (+/- 0.000013) for *rad51-4* -/- X and Y respectively.



**Figure 3-12:  $IC_{50}$  values for methyl methanesulfonate (MMS) for *RAD51-4* mutant cell lines**

Resazurin was used as an indicator of growth and  $IC_{50}$  values were calculated for each cell line. Wildtype 3174.2 (wt), two independent *RAD51-4* heterozygous mutants (X+/-, Y+/-), two independent *rad51-4* homozygous mutants (X-/-, Y-/-) and two independent *RAD51-4* re-expressor cell lines (X-/-+/, Y-/-+/) were treated with serial dilutions of MMS. The graph shows the mean  $IC_{50}$  values determined from the triplicate data, with the 95 % confidence intervals indicated by the error bars.

The reduction of the  $IC_{50}$  value of the *rad51-4* -/- mutants compared to the wt cells was found to be significantly different (see Table 3-5) as judged by the results of the unpaired, two-tailed Student's *t*-test comparisons of the means of the triplicate results when plotted by nonlinear regression on a sigmoid dose response graph. In contrast, there was no

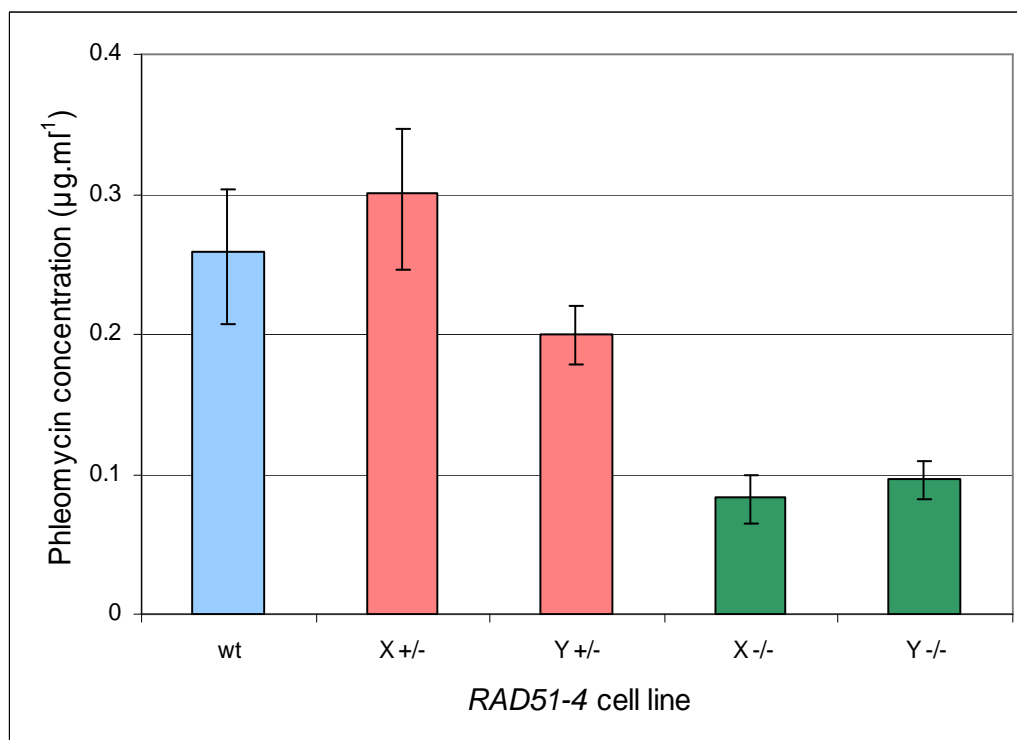
evidence that the *RAD51-4* +/- mutants or the *RAD51-4* -/-/+ cells displayed this impairment (Appendix 3, Table 4).

**Table 3-5: Statistical analysis of the IC<sub>50</sub> values for methyl methanesulfonate (MMS) of the *RAD51-4* mutants.**

Student's *t*-tests were carried out to compare the mean of the triplicate IC<sub>50</sub> values for the cell lines against the wildtype (wt), a significant difference between the results is shown by a *p*-value of  $p \leq 0.05$ , not significantly different results are shown as "-". [NA is not applicable.] Each cell line is indicated as follows: wildtype strain 3174.2 (wt), two independent *RAD51-4* heterozygous mutants (X+/-, Y+/-), *rad51-4* homozygous mutants (X-/-, Y-/-) and *RAD51-4* re-expressors (X-/-/+, Y-/-/+).

	wt	X+/-	Y+/-	X-/-	Y-/-	X-/-/+	Y-/-/+
wt	NA	-	-	0.0120	0.0387	-	-

Figure 3-13 shows the mean IC<sub>50</sub> values for phleomycin of the wildtype, *RAD51-4* +/-, and *rad51-4* -/- cell lines. For this DNA damaging compound, the wildtype cells' mean IC<sub>50</sub> value was 0.26  $\mu\text{g}.\text{ml}^{-1}$  (+/- 0.05), which was 2.3 fold greater than the *rad51-4* -/- mutants: X-/- mean IC<sub>50</sub> value was 0.08  $\mu\text{g}.\text{ml}^{-1}$  (+/- 0.2) and Y-/- was 0.09  $\mu\text{g}.\text{ml}^{-1}$  (+/- 0.01).



**Figure 3-13: IC<sub>50</sub> values for phleomycin in *RAD51-4* mutant cell lines.**

Resazurin was used as an indicator of growth and IC<sub>50</sub> values were calculated for each cell line. Wildtype 3174.2 (wt), two independent *RAD51-4* heterozygous mutants (X+/-, Y+/-), *RAD51-4* homozygous mutants (X-/-, Y-/-) and *RAD51-4* re-expressor cell lines (X-/-/+, Y-/-/+) were treated with serial dilutions of phleomycin. The graph shows the mean IC<sub>50</sub> derived from the triplicate data, with the 95 % confidence intervals indicated by the error bars.

The results of Student's *t*-test analysis of these data are shown in Table 3-6. The reduction of the IC<sub>50</sub> value for phleomycin treatment of the *rad51-4* -/- mutant relative to wildtype



was significant. The complete statistical analysis of all pair-wise comparisons can be found in Appendix 3, Table 5.

**Table 3-6: Statistical analysis of the IC<sub>50</sub> values for phleomycin of the *RAD51-4* mutants.**

Student's *t*-tests were carried out to compare the mean of the triplicate IC<sub>50</sub> values for the cell lines against the wildtype, a significant difference between the results is shown by a *p*-value of  $p \leq 0.05$  and not significantly different results are shown as “-”. [NA is not applicable.] Each cell line is initiated as follows: wildtype strain 3174.2 (wt), two independent *RAD51-4* heterozygous mutants (X+/-, Y+/-), and *rad51-4* homozygous mutants (X-/-, Y-/-).

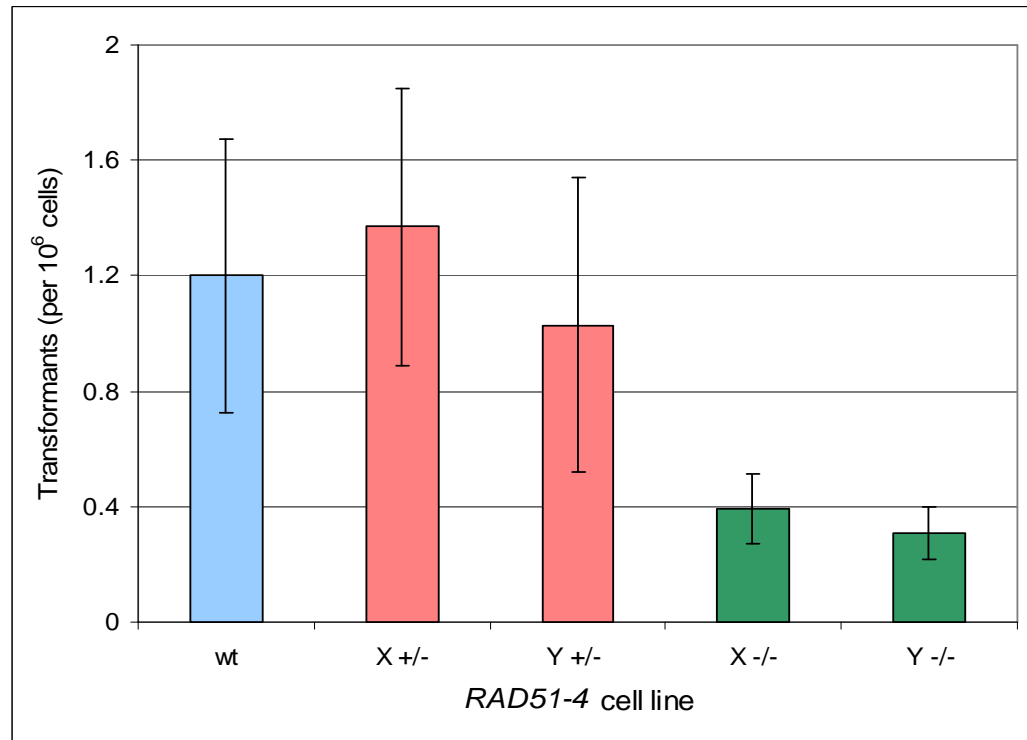
	wt	X+/-	Y+/-	X-/-	Y-/-
wt	NA	-	-	0.0280	0.0499

These data confirm the findings of the clonal survival analysis and confirms that the *rad51-4* -/- mutants are more sensitive to two different DNA-damaging agents, due to the lack of functional *RAD51-4* in the cell. The re-expressor restores most or all of the resistance to the MMS, in both assays, although this could not be confirmed for the sensitivity of phleomycin (Section 3.3).

### 3.4.3 Analysis of the recombination efficiency of *RAD51-4* mutant cell lines

The analysis of the *rad51-4* -/- mutants ability to undergo homologous recombination was analysed next. To do this, the cells' ability to take up DNA and integrate into the genome was quantified using a transformation efficiency test. To assay for transformation efficiency, a linearised DNA construct containing the ORF of phleomycin resistance, flanked by  $\beta\alpha$  and  $\alpha\beta$  intergenic regions, was transformed into the wildtype and each of the mutant cell lines (see Materials and Methods, Section 2.8.3). If homologous recombination can occur, the targeting regions of  $\beta\alpha$  and  $\alpha\beta$  intergenic flanks allow the resistance cassette to integrate into the  $\alpha$  tubulin array, and as result the cells become resistant to phleomycin. Each cell line at a total volume of  $5 \times 10^7$  cells was transformed by electroporation with 5  $\mu\text{g}$  of linearised DNA. After transformation, the cells were allowed to recover for 24 hours prior to phleomycin selection. The cells were incubated for 14 days at 37 °C on 2.0  $\mu\text{g}.\text{ml}^{-1}$  of phleomycin and the number of wells with transformants was scored. The transformation efficiency was described as the number of wells with growth per  $10^6$  cells plated out.

Figure 3-14 shows the results of this analysis and indicates that the *rad51-4* -/- mutants have a decreased ability to undergo transformation, since the transformation efficiency of wildtype cells was  $1.2 \times 10^{-6}$  (+/-  $4.75 \times 10^{-7}$ ) compared with  $3.92 \times 10^{-7}$  (+/-  $1.2 \times 10^{-7}$ ) and  $3.8 \times 10^{-7}$  (+/-  $8.92 \times 10^{-8}$ ) for the two *rad51-4* -/- mutants, X-/- and Y-/- respectively.



**Figure 3-14: Transformation efficiency assay for the *RAD51-4* mutant cell lines.**

To assay for recombination rate, a transformation efficiency assay was used. Each cell line is indicated as follows: Wildtype strain 3174.2 (wt), two independent *RAD51-4* heterozygous mutants (X+/-, Y+/-), and two independent *rad51-4* homozygous mutants (X-/-, Y-/-). The values plotted are mean numbers of phleomycin resistant transformants per 10<sup>6</sup> cells transformed. These experiments were done in triplicate with the same linearised DNA, and error bars of the 95 % confidence intervals are indicated.

In contrast, disruption of any one *RAD51-4* allele did not affect the cells ability to undergo transformation, since the *RAD51-4* +/- cells had comparable transformation rates to that of wt cells. The decreased transformation rates of the *rad51-4* -/- cells were confirmed to be significantly different, using unpaired, two-tailed Student's *t*-test, and are shown in Table 3-7. The complete statistical analysis of all pair-wise comparisons is shown in Appendix 3, Table 6.

**Table 3-7: Statistical analysis of the transformation efficiency of *RAD51-4* mutant cell lines.**

Unpaired, two-tailed Student's *t*-tests were carried out to compare the mean of the triplicate transformation efficiency results for the wildtype cell (wt) relative to the other cell lines. A significant difference between the results is shown by a *p*-value of  $p \leq 0.05$ . Not significantly different results are shown as "-". NA is not applicable. Each cell line is indicated as follows: wildtype strain 3174.2 (wt), two independent *RAD51-4* heterozygous mutants (X+/-, Y+/-), and *rad51-4* homozygous mutants (X-/-, Y-/-).

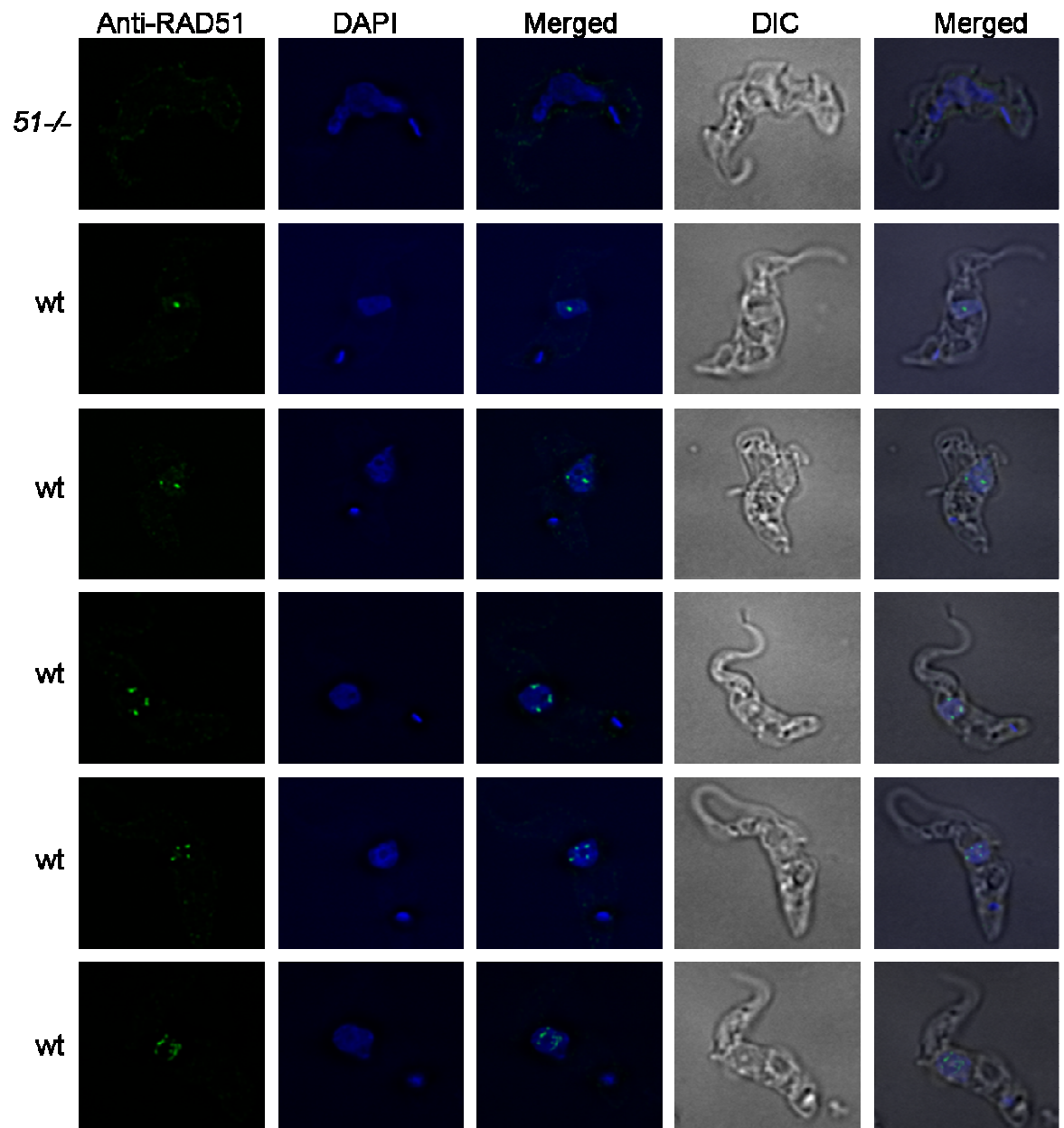
	wt	X+/-	Y+/-	X-/-	Y-/-
wt	NA	-	-	0.0089	0.0047

This analysis has indirectly shown that the *rad51-4* -/- mutants have a decrease in their ability to undergo homologous recombination compared to the wildtype and *RAD51-4* +/- mutants. This phenotype is highly similar to that observed in the *rad51* -/-, *rad51-3* -/-,

*rad51-5* *-/-* and *brca2* *-/-* cell lines, where there was also a decrease in transformation efficiency, but notably there was no obliteration of the cell recombination capacity (Hartley, 2008; McCulloch & Barry, 1999; Proudfoot & McCulloch, 2005a). In all previous cases, the residual recombination occurred by homologous recombination, indicating a number of pathways operate. Similar analysis was not conducted here, however.

### **3.4.4 Analysis of RAD51 foci formation in the *RAD51-4* mutant cell lines**

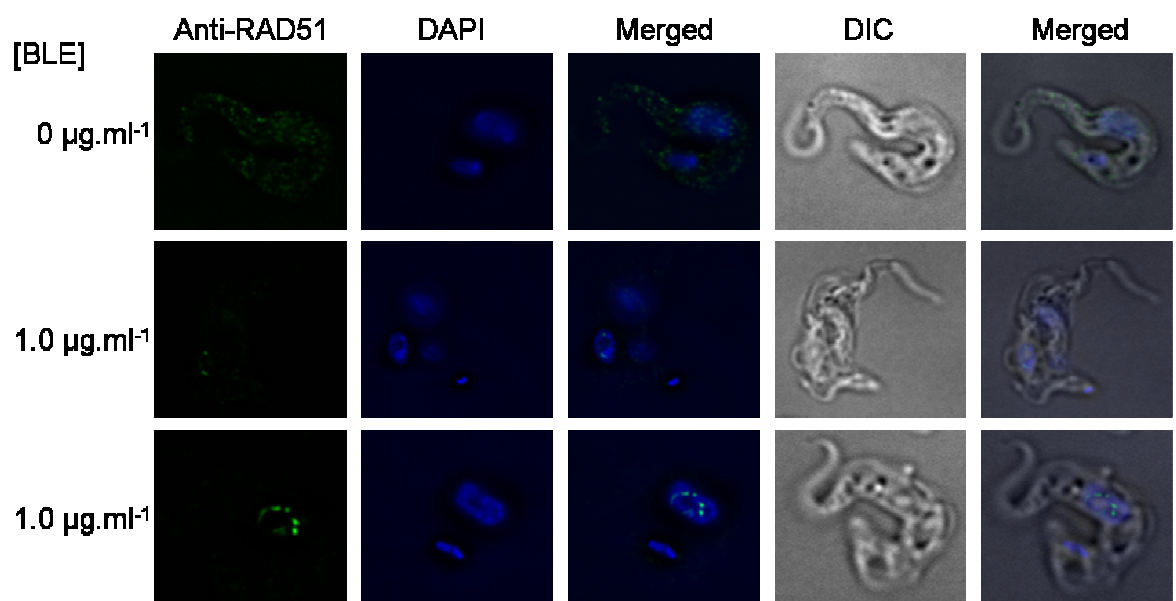
In eukaryotic cells, it has been shown that RAD51 forms discrete foci in sub-nuclear complexes contributing to the repair of damaged DNA including stalled/collapsed replication forks, and during meiosis (Bishop, 1994; Haaf *et al.*, 1995; Haaf *et al.*, 1999; Lisby *et al.*, 2004; Plug *et al.*, 1996; Scully *et al.*, 1997; Tarsounas *et al.*, 2004a; Ward *et al.*, 2007). It has also been shown that RAD51 foci formation can be induced by DNA damaging agents such as UV radiation (Raderschall *et al.*, 1999; Tarsounas *et al.*, 2004a; Tashiro *et al.*, 2000). RAD51 foci are most probably formed at the sites of DNA damage (Sugawara *et al.*, 2003). It has been previously shown that DNA damage by phleomycin treatment of *T. brucei* causes a similar re-localisation of RAD51 into distinct foci, and that RAD51-3, RAD51-5 and BRCA2 all influence the formation of these complexes (Glover *et al.*, 2008; Hartley, 2008; Proudfoot & McCulloch, 2005a). The number of RAD51 foci per cell varies, including in *T. brucei*, presumably in correlation with the amount of DNA-damage induced in the cell, although this has not been quantitatively analysed in *T. brucei*. Figure 3-15 shows examples of these foci following phleomycin treatment of wildtype bloodstream stage cells in which one to five distinct foci are seen. The Figure 3-15 also shows treated *rad51* *-/-* cell lines, which do not form foci due to RAD51 being absent.



**Figure 3-15: RAD51 sub-nuclear foci in *T. brucei*.**

The first row of pictures shows an example of a phleomycin treated *rad51*<sup>-/-</sup> cell: no distinct RAD51 foci are detected. The other rows are examples of the number of RAD51 foci formed in treated wildtype BSF *T. brucei* cells, (1, 2, 3, 4 and many). RAD51 was imaged by indirect immuno-fluorescence (FITC) (first column), the DNA was stained with DAPI (second column), FITC and DAPI were merged (third column), the cells are visualised by DIC (fourth column) and all the images were merged (fifth column).

To examine the role of RAD51-4 in RAD51 foci formation immuno-fluorescence was performed on the *rad51-4*  $-/-$  mutant cell lines (see Materials & Methods, Section 2.9) using rabbit anti-RAD51 anti-serum at a dilution of 1:500 and detected using goat anti rabbit anti-serum conjugated with SFX (fluorescein, succinimidyl ester: Molecular Probes). This was performed for wildtype, and each independent *RAD51-4*  $+/-$  and *rad51-4*  $-/-$  mutants, all of which had been grown on varying concentrations of phleomycin for 18 hours. All cell lines in the absence of phleomycin treatment showed minimal RAD51 foci formation, with only 1-3 % having detectable foci, as no DNA damage was induced; an example of this is shown in Figure 3-16. At varying concentrations of phleomycin, from  $0.25 \mu\text{g.ml}^{-1}$  to  $2.0 \mu\text{g.ml}^{-1}$ , 80 % - 100 % of wt cells produced RAD51 foci (Table 3-8). The *RAD51-4*  $+/-$  cell lines also had a similar pattern to wildtype, with 65 % - 93 % of cells forming RAD51 foci.



**Figure 3-16: RAD51 sub-nuclear foci in *rad51-4*  $-/-$  mutant cell lines.**

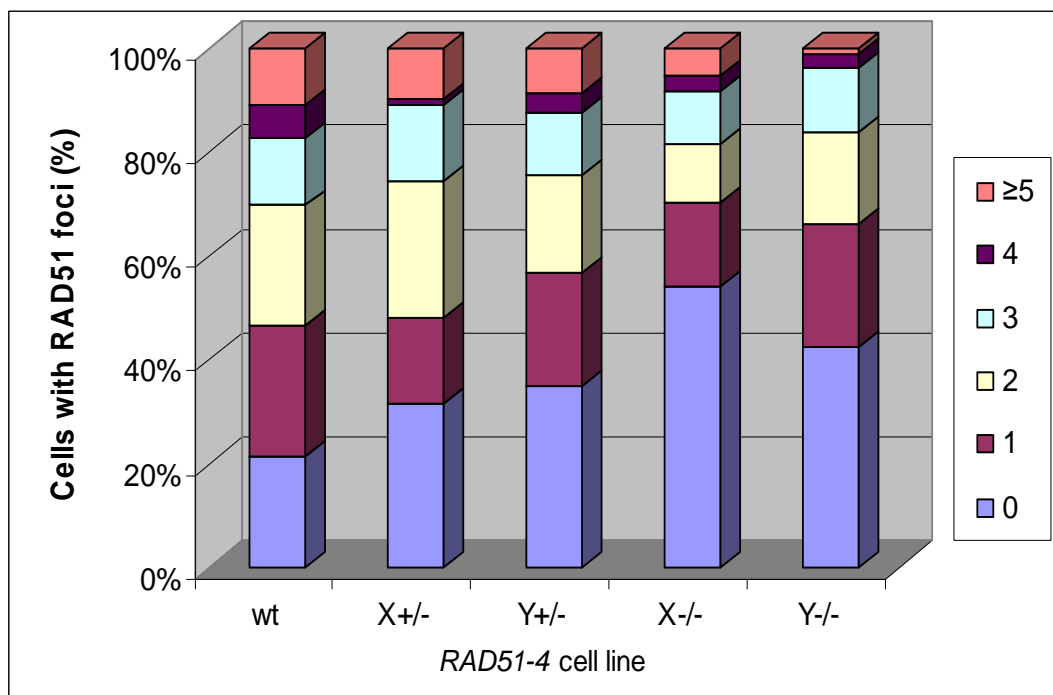
The first row of pictures shows an example of a non-treated *rad51-4*  $-/-$  cell: no distinct RAD51 foci are detected. The second row is an example of a phleomycin treated *rad51-4*  $-/-$  cell where no foci are present. The last row is an example of a phleomycin treated *rad51-4*  $-/-$  cell where five RAD51 foci are present in the nucleus. RAD51 was imaged by indirect immuno-fluorescence (FITC) (first column), the DNA was stained with DAPI (second column), these two were merged (third column), the cells are visualised by DIC (fourth column) and all the images were merged (fifth column).

**Table 3-8: Percentage of *rad51-4*  $-/-$  mutant cells with a specific number of RAD51 foci.**

The percentage of cells with a specific number of sub-nuclear RAD51 foci formed (0, 1, 2, 3, 4,  $\geq 5$ ) when the cell lines were treated with increasing concentrations of phleomycin,  $0.25 \mu\text{g.ml}^{-1}$ ,  $1.0 \mu\text{g.ml}^{-1}$  and  $2.0 \mu\text{g.ml}^{-1}$  is shown. The total indicates the number of cells counted. [Wildtype strain 3174.2 is indicated by wt, two independent *RAD51-4* heterozygous mutants are indicated by X $+/$ - and Y $+/$ -, *rad51-4* homozygous mutants are indicated by X $-/-$  and Y $-/-$ .

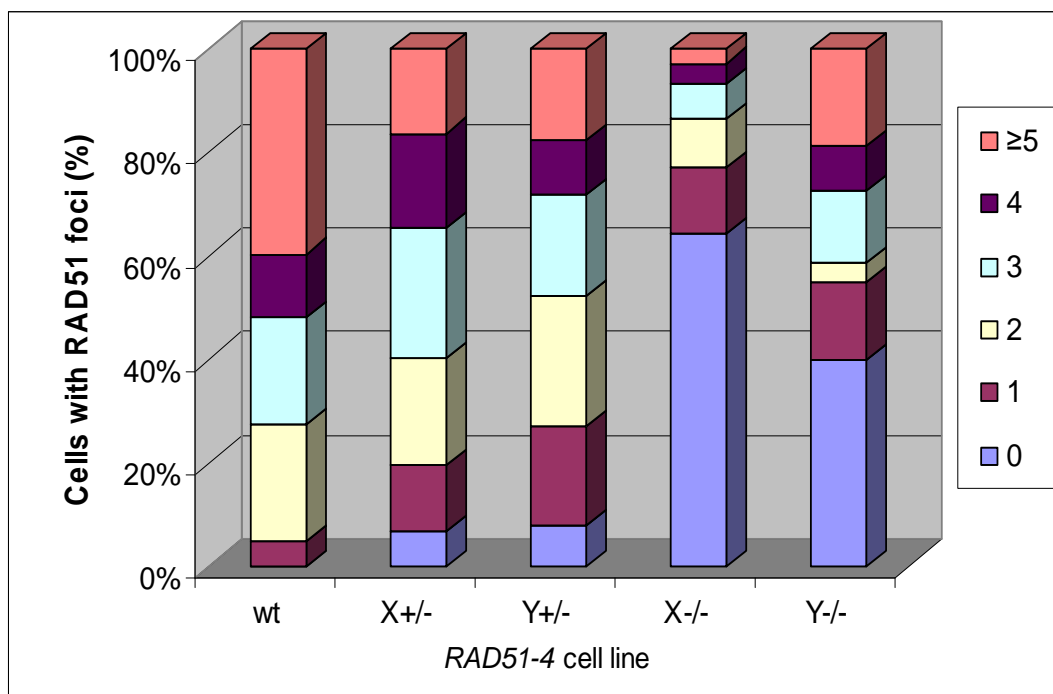
	Concentration of phleomycin	Number of RAD51 foci						Cells Counted
		0	1	2	3	4	$\geq 5$	
wt	$0.25 \mu\text{g.ml}^{-1}$	32.0	43.0	14.0	9.0	1.0	1.0	100
	$1.0 \mu\text{g.ml}^{-1}$	21.3	25.3	23.1	12.9	6.7	10.7	225
	$2.0 \mu\text{g.ml}^{-1}$	0.0	4.7	22.6	20.8	12.3	39.6	106
X $+/$ -	$1.0 \mu\text{g.ml}^{-1}$	31.4	16.7	26.5	14.7	1.0	9.8	100
	$2.0 \mu\text{g.ml}^{-1}$	6.7	12.9	20.6	25.4	18.2	16.3	209
Y $+/$ -	$1.0 \mu\text{g.ml}^{-1}$	34.9	21.7	18.9	12.3	3.8	8.5	106
	$2.0 \mu\text{g.ml}^{-1}$	7.9	19.0	25.4	19.6	10.6	17.5	189
X $-/-$	$0.25 \mu\text{g.ml}^{-1}$	61.3	19.1	11.9	3.4	2.6	1.7	235
	$0.5 \mu\text{g.ml}^{-1}$	48.5	30.0	14.5	3.5	2.5	1.0	200
	$1.0 \mu\text{g.ml}^{-1}$	54.1	16.3	11.5	10.0	2.9	5.3	209
	$2.0 \mu\text{g.ml}^{-1}$	64.2	12.8	9.5	6.8	4.1	2.7	148
Y $-/-$	$0.25 \mu\text{g.ml}^{-1}$	87.4	9.7	1.9	0.0	0.0	1.0	103
	$0.5 \mu\text{g.ml}^{-1}$	37.4	29.1	19.7	8.9	1.5	3.4	203
	$1.0 \mu\text{g.ml}^{-1}$	42.5	23.6	17.9	12.3	2.8	0.9	108
	$2.0 \mu\text{g.ml}^{-1}$	40.0	15.0	3.8	13.8	8.8	18.8	80

The *rad51-4*  $-/-$  mutants appear to be less able to form RAD51 foci. At a phleomycin concentration of  $1.0 \mu\text{g.ml}^{-1}$ , 21 % of wildtype cells did not produce detectable RAD51 foci (see Table 3-8 and Figure 3-17), while in the two independent *rad51-4*  $-/-$  cell lines, 54 % and 42 % of the total counted do not produce RAD51 foci, for X $-/-$  and Y $-/-$  respectively. This difference was more marked on increasing the phleomycin concentration to  $2.0 \mu\text{g.ml}^{-1}$  (see Figure 3-18). In these conditions, all of the wildtype cells produce RAD51 foci, while 64 % and 40 % of the *rad51-4*  $-/-$  mutants cells (X $-/-$  and Y $-/-$  respectively) do not produce RAD51 foci. Similarly, at the lower concentration of  $0.25 \mu\text{g.ml}^{-1}$ , 87 % and 61 % of the *rad51-4*  $-/-$  cells displayed no foci, compared to only 32 % of the wildtype cells. These data indicate the lack of RAD51-4 in the cell causes a defect in RAD51 re-localisation to or within the nucleus when the cells are treated with  $2.0 \mu\text{g.ml}^{-1}$  of phleomycin. It is noticeable however, that discernible RAD51 foci do occur, indicating this re-localisation is not abolished.



**Figure 3-17: Percentage of *RAD51-4* mutant cell lines with a specific number of the RAD51 foci when treated with 1.0  $\mu\text{g.ml}^{-1}$  of phleomycin.**

The percentage of cells with a specific number of the sub-nuclear RAD51 foci formed (0, 1, 2, 3, 4,  $\geq 5$ ) is shown when the cell lines were treated with 1.0  $\mu\text{g.ml}^{-1}$  of phleomycin. The cell lines are labelled as follows: wildtype strain 3174.2 (wt), the two independent *RAD51-4* heterozygous mutants (X+/-, Y+/-), and the two independent *rad51-4* homozygous mutants (X-/-, Y-/-).



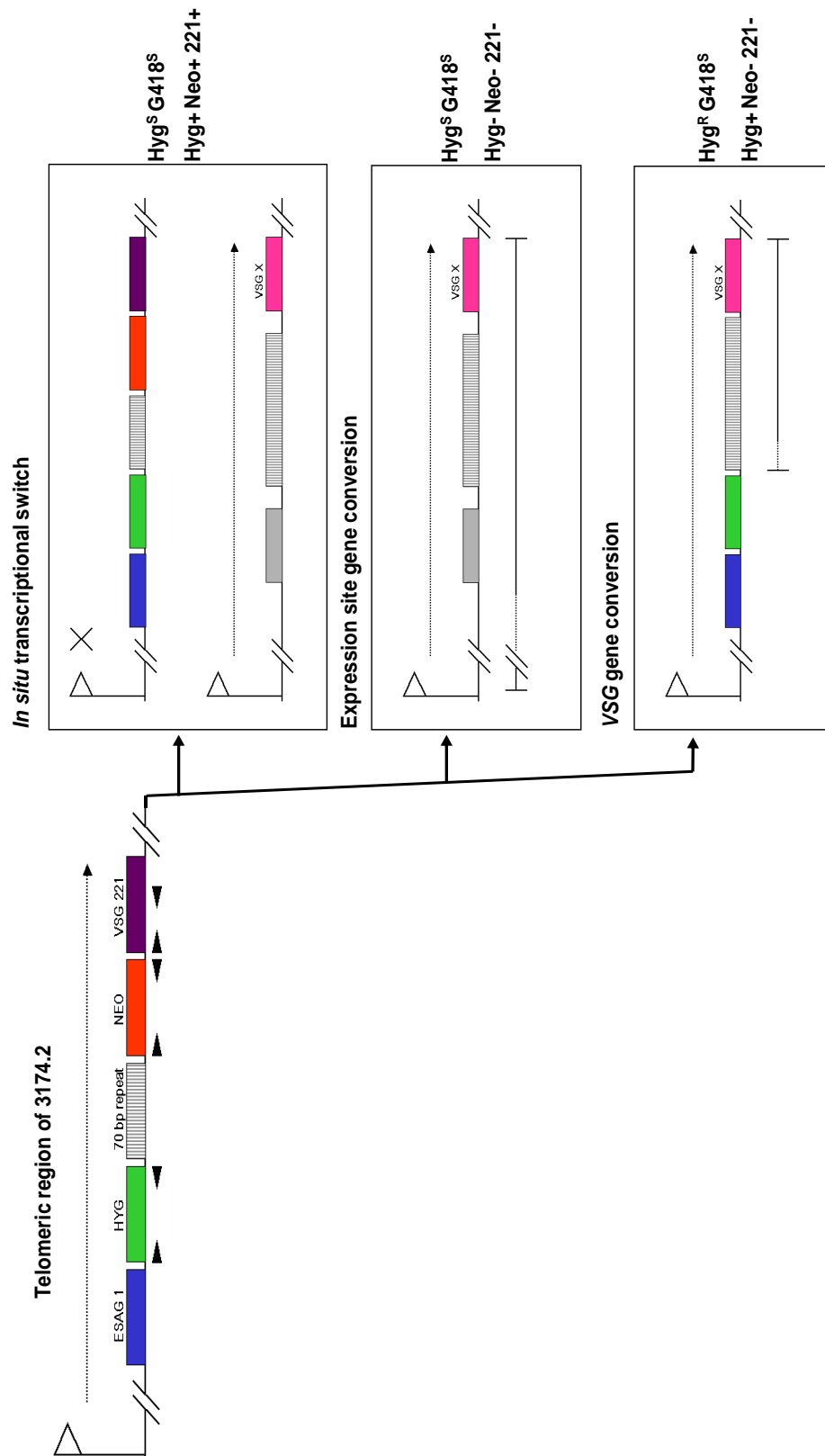
**Figure 3-18: Percentage of *RAD51-4* mutant cells with a specific number of the RAD51 foci when treated with 2.0  $\mu\text{g.ml}^{-1}$  of phleomycin.**

The percentage of cells with a specific number of the sub-nuclear RAD51 foci formed (0, 1, 2, 3, 4,  $\geq 5$ ) is shown when the cell lines were treated with 2.0  $\mu\text{g.ml}^{-1}$  of phleomycin. The cell lines are labelled as follows: wildtype strain 3174.2 (wt), the two independent *RAD51-4* heterozygous mutants (X+/-, Y+/-), and the two independent *rad51-4* homozygous mutants (X-/-, Y-/-).

### 3.4.5 Analysis of VSG switching of mutant cell lines

The 3174.2 strain of *T. brucei* was used to generate the *rad51-4* mutants in bloodstream stage cells (see Materials and Methods, Section 2.1.1). This strain has been described before and can be used to analyse the VSG switching frequency and mechanisms (Bell & McCulloch, 2003; Conway *et al.*, 2002a; McCulloch *et al.*, 1997; Proudfoot & McCulloch, 2005a; Robinson *et al.*, 2002). As previously described, *rad51*, *rad51-3* and *brca2* mutations cause a decrease of VSG switching frequency (Hartley, 2008; McCulloch & Barry, 1999; Proudfoot & McCulloch, 2005a). In contrast, there is no evidence that RAD51-5 acted in this process which may suggest some specialised role of the RAD51 paralogue (Proudfoot & McCulloch, 2005b). The 3174.2 strain was also used to determine if the absence of RAD51-4 affects antigenic variation by analysing the VSG switching frequency and mechanism. This cell line has a modified active VSG expression site (ES) as it contains two antibiotic resistance cassettes inserted upstream of the telomeric VSG, named *VSG221*. The frequency of VSG switching was measured by an *in vivo* assay involving mice immunised against *VSG221* (see Materials and Methods, Section 2.8.4). The resistance cassettes, hygromycin (Hyg, *HYG*) and neomycin (G418, *NEO*), ensure expression of the *VSG221* when the cells are grown on the antibiotic *in vitro*, allowing injection of and immunisation of the mice against *VSG221*-expressing cells. Once the mice have recovered from challenge and cure of *T. brucei* expressing *VSG221*, they were re-infected with the cell lines grown for nine generations *in vitro* without the selection of Hyg and G418, allowing VSG switching to occur and switched variants to arise. By measuring the number of *T. brucei* cell survivors after re-infection of the mice, the rate of switching was calculated. The switching mechanisms used in cloned variants that were recovered were also determined by identification of the resistance markers and their expression in the ES (see Figure 3-19).





**Figure 3-19: Use of the transgenic *T. brucei* strain 3174.2 allows characterisation of the VSG switching events.**

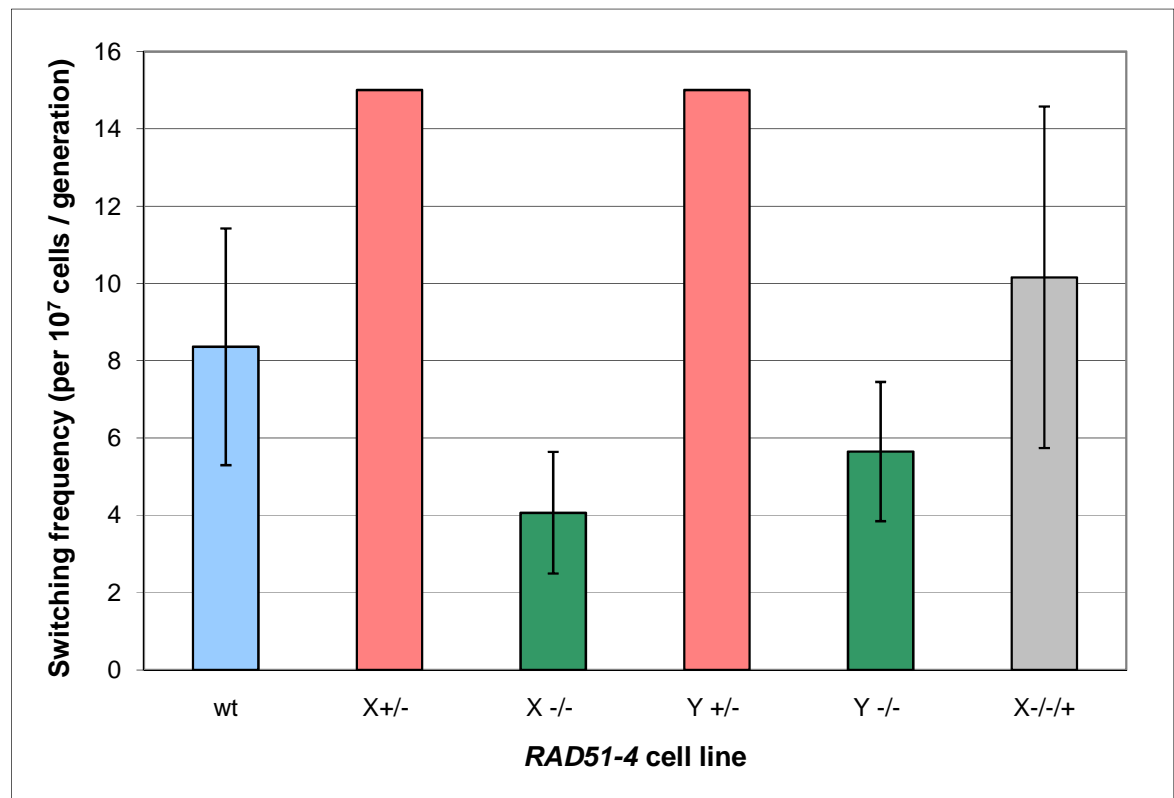
On the left, the diagram shows the active expression site (ES) with *VSG221* (purple box) with the two antibiotic cassettes, hygromycin (*HYG*; green box) and neomycin (*NEO*; red box) which are separated by the 70 bp repeats (grey box). In this active ES, there is also expression site associated gene 1 (ESAG1; blue box). The primer binding sites are denoted by the black arrows. On the right of the diagram, the three possible switching events, *in situ* transcriptional switch, expression site gene conversion/loss or VSG gene conversion are shown. G418 and Hyg resistance (R) or sensitivity (S) along with the PCR product presence (+) or absence (-) are indicated. Taken from Hartley, C thesis and adapted from McCulloch *et al.*, 1997.

The switching frequency was calculated by counting the number of wells in a 96 well plate that contained growing trypanosome populations up to 14 days after the cells were recovered from the re-challenged immune mice. A selection of recovered, clonal populations was then grown on Hyg and G418 to ascertain if the genes were expressed and the genomic DNA from each population was extracted to allow PCR-amplification to test for the presence or absence of the *HYG*, *NEO* and *VSG221* ORFs (see Materials and Methods, Section 2.4.1). This allowed the *VSG* switching mechanism to be deduced by PCR-amplification and drug resistances/sensitivities. The three possible switching events are described in Figure 3-19. During an *in situ* transcriptional switch the *VSG221* expression site is inactivated. There is no loss or replacement of the genes, and an inactive expression site containing a different *VSG*, *VSG(X)* is transcriptionally activated. As a result, any cells that have undergone this type of switching mechanism are sensitive to both drugs, (Hyg and G418 sensitive: Hyg<sup>s</sup> and G418<sup>s</sup> respectively). As there is no loss of the genes, PCR of the expression site shows that the *HYG*, *NEO* and *VSG221* gene are present (Hyg+, Neo+, VSG+). In an expression site gene conversion event, a different *VSG(X)* replaces all or a significant part of the *VSG221* expression site. As a result, any cells that have undergone this type of switching mechanism are Hyg<sup>s</sup> and G418<sup>s</sup> and as there is loss of the genes the PCR-amplified products of the expression site markers shows that *HYG*, *NEO* and *VSG221* gene are not present (Hyg-, Neo-, VSG-). In a *VSG* gene conversion event, a different *VSG(X)* replaces the region from the 70 bp repeats to the end of the *VSG221*. As a result, any cells that have undergone this type of switching mechanism are Hyg resistant because it is still actively transcribed (Hyg<sup>r</sup>) and G418 sensitive as the neomycin resistance cassette is in between the 70 bp repeats and the *VSG221* gene (G418<sup>s</sup>); PCR-amplification products of the expression site markers shows that *HYG* is present, and *NEO* and *VSG221* genes are not present (Hyg+, Neo-, VSG-).

### 3.4.6 Analysis of *VSG* switching of *RAD51-4* mutant cell lines

The switching frequencies and mechanisms for the *RAD51-4* cell lines were analysed as described in Section 3.4.5, and Materials and Methods, Section 2.8.4. Briefly, each cell line at a cell density of  $4 \times 10^7$  cells, was injected into the previously immunised mice. *VSG* switching frequency was calculated as follows: The isolated and cloned surviving switched *T. brucei* cells were counted over 96-well plates. This number was multiplied by 2.5, as the recovered blood from the mouse (plated out 0.4 ml) was only 40 % of the total approximate blood volume of the animal. This number was then divided by the number of doubling times that had occurred during the 24 hour period following injection into the

mouse. Finally this value was divided by the number of cells grown *in vitro* that had been injected into the VSG221-immunised mice. For each VSG switch frequency, a minimum of three independent experiments were performed.



**Figure 3-20: The mean VSG switching frequencies of the *RAD51-4* cell lines.**

Wildtype cells (wt), the two independent *RAD51-4* heterozygote cell lines (X+/-, Y+/-), the two independent *rad51-4* homozygote cell lines (X-/-, Y-/-) and one *RAD51-4* re-expressor cell line (X-/-/+) was injected into a minimum of three mice to analyse whether or not the *rad51-4* -/- mutant had an effect on VSG switching frequencies. The graph shows the mean VSG switching frequencies and is described as events per 10<sup>7</sup> cells per generation. The 95 % confidence intervals are indicated as error bars.

Figure 3-20 shows the mean VSG switching frequencies of the cell lines examined. The wildtype cells' VSG switching frequency was 8.36 (+/- 3.06) switched variants per 10<sup>7</sup> cells per generation, while the *rad51-4* -/- mutants (X-/-, Y-/-) had a VSG switching frequency of 4.06 (+/- 1.57) and 5.64 (+/- 1.8) switched variants per 10<sup>7</sup> cells per generation respectively. Each *RAD51-4* heterozygote mutant (X+/-, Y+/-) had a high recovery rate, with growth in all 96 wells for all three mice. It is unknown why this was the case. *RAD51-4* X+/- had a VSG switching frequency of 15 switched variants per 10<sup>7</sup> cells per generation compared the significantly different *rad51-4* X-/- . The same was true for the other heterozygote independent clone (Y+/-) which also had a VSG switching frequency of 15 switched variants per 10<sup>7</sup> cells per generation, and was significantly different to the *rad51-4* Y-/- . This analysis showed that *rad51-4* mutants had an impaired ability to undergo VSG switching as the VSG switching frequencies of the *rad51-4* -/- were greatly reduced compared with the cell lines from which they were derived: *RAD51-4* +/- mutants.

This is not unequivocal evidence that RAD51-4 is involved in VSG switching, as there was no significant difference between both independent homozygous mutants and the wildtype cells.

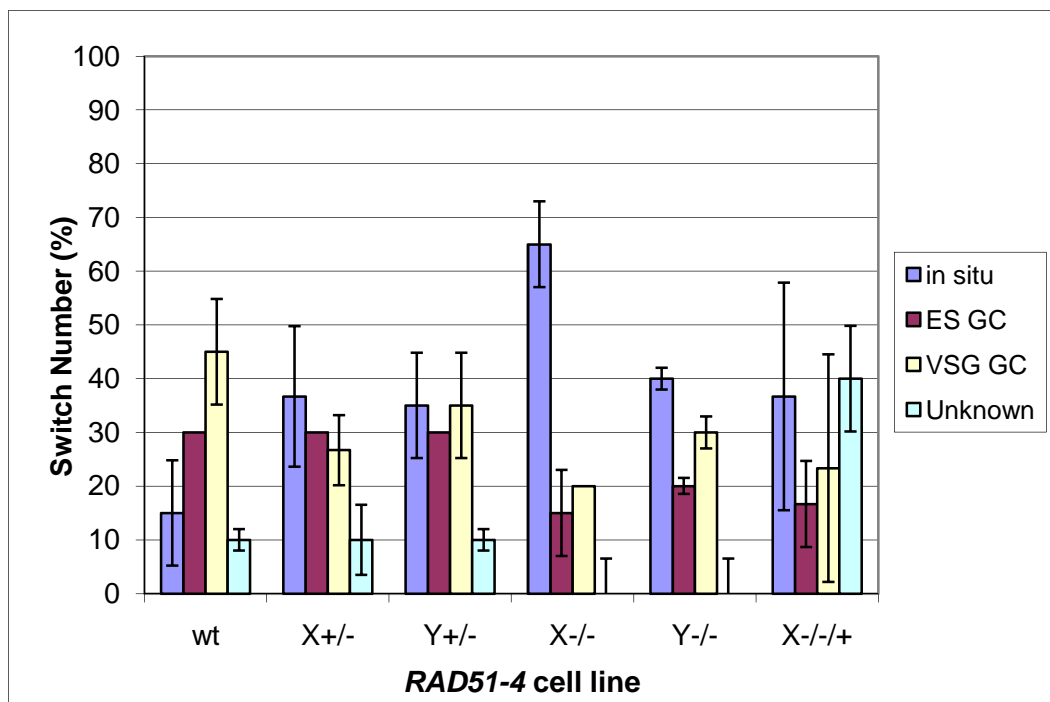
Table 3-9 shows that there was a significant difference between the data for the wildtype compared to the X homozygous mutant and not the Y homozygous mutant, as well as a significant difference between the wildtype and one of the *RAD51-4* +/- heterozygous mutants (X+/-). This is mostly likely due to the high variability and nature of the *in vivo* experiment. Unpaired two-tailed Student's *t*-tests were carried out on all the data, and are shown in Appendix 3, Table 7.

**Table 3-9: Statistical analysis of the VSG switching frequencies of *RAD51-4* mutant cell lines.**

Unpaired, two-tailed Student's *t*-tests were carried out to compare the mean of the VSG switching frequencies for the cell lines against the wildtype. A significant difference is shown by a *p*-value of  $p \leq 0.05$ . Not significantly different results are shown as "-". [NA is not applicable.] Wildtype strain 3174.2 (wt), two independent *RAD51-4* heterozygous mutants (X+/-, Y+/-), two independent *rad51-4* homozygous mutants (X-/-, Y-/-) and the *RAD51-4* re-expressor (X-/-/+) are shown.

	wt	X+/-	Y+/-	X-/-	Y-/-	X-/-/+
wt	NA	0.0107	-	0.0285	-	-

A selection of recovered clones from each cell line (10 from each mouse; a total of 30 for each cell line) were analysed to differentiate the switching mechanisms that occurred, described in Section 3.4.5 and Materials and Methods, and Section 2.8.4. As RAD51-4 act in homologous recombination and DNA repair it might be suggested that the *rad51-4* -/- would not undergo gene conversion events to switch their VSG as this mechanism relies on homologous recombination. Figure 3-21 shows the switching mechanisms in the *rad51-4* -/- mutants. This data does not fit the prediction, and shows that the *rad51-4* -/- mutant, clone Y-/-, can undergo expression site and VSG gene conversion. In contrast, there is an observable difference between the wildtype *in situ* switch compared with the *rad51-4* X-/- mutant clone. It was observed that 15 % (+/- 9.7 %) of wildtype cells used an *in situ* switch whilst 65 % (+/- 8.0 %) of the *rad51-4* X-/- had switched their VSG by an *in situ* switch. This clear difference was not observed in the other *rad51-4* Y-/- . It should also be noted that both *rad51-4* -/- mutants were still able to undergo expression site and VSG gene conversion. The drug resistance and sensitivities along with the PCR-banding pattern of a number of switchers, which can not define by the three mechanism described in Figure 3-19, are shown as "unknown". These unknown mechanisms may be the result of single-strand annealing pathway as a way of switching the VSG gene in the expression site.



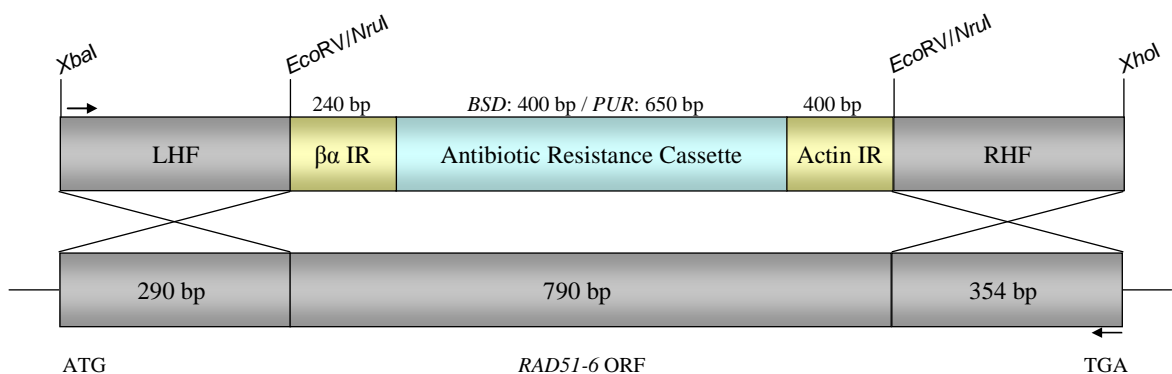
**Figure 3-21: Analysis of the switching mechanisms of the *RAD51-4* mutant cell lines.**

The graph shows the switching mechanism of *RAD51-4* mutant cell lines as a percentage of the total analysed. Wildtype strain 3174.2 (wt), two independent *RAD51-4* heterozygous mutants (X+/-, Y+/-), two independent *rad51-4* homozygous mutants (X-/-, Y-/-) and the *RAD51-4* re-expressor (X-/-/+) are shown. The assay allowed differentiation of *in situ* transcriptional switching (*in situ*), expression site gene conversion (ES GC) and VSG gene conversion (VSG GC). Switch reactions that could not be assigned in these categories are indicated as “unknown”. The graph shows the means of triplicate data, with the 95 % confidence intervals indicated by the error bars.

## 3.5 Generation of the *RAD51-6* gene disruption mutants

### 3.5.1 Design of the *RAD51-6* knockout constructs

The constructs for generating *RAD51-6* disruption mutants were of the same design as the *RAD51-4* knockout constructs (Section 3.2). In this case, recombination on the LHF and RHF targeting regions of the *RAD51-6* ORF, resulted in the deletion of 790 bp of the ORF which was replaced with the *BSD* and *PUR* antibiotic cassettes (see Figure 3-22). The nucleotide sequence of the *RAD51-6* ORF and the primer binding sites are shown in Appendix 2. The LHF of *RAD51-6* was PCR-amplified using primers *RECA2* StartORF For and *RECA2* StartORF Rev. The RHF of *RAD51-6* was PCR-amplified using primers *RECA2* EndORF For and *RECA2* EndORF Rev, see Materials and Methods, Section 2.4.1, Table 2-1. These PCR-amplified products, of 290 bp and 354 bp respectively, were cloned into *Xba*I and *Xho*I digested pBluescript KS (Stratagene) using the same strategy as for the *RAD51-4* knockout constructs. The *BSD* and *PUR* antibiotic resistance cassettes were ligated between the LHF and RHF of *RAD51-6* having previously been PCR-amplified and cloned into pCR®2.1 (see Section 2.4.7). Once the insertion and correct orientation of the resistance cassettes between the *RAD51-6* LHF and RHF were confirmed by restriction mapping and sequencing, the resulting constructs were called  $\Delta$ *RAD51-6::BSD* and  $\Delta$ *RAD51-6::PUR*.



**Figure 3-22: *RAD51-6* gene disruption construct and gene disruption strategy.**

LHF: Left hand targeting flank which has homology to the 5' region of the *RAD51-6* ORF. RHF: Right hand targeting flank which had homology to the 3' region of the *RAD51-6* ORF. The 5' targeting flank (LHF) was approximately 290 bp and the 3' targeting flank (RHF) was 354 bp. The antibiotic cassettes have  $\beta\alpha$  tubulin intergenic region ( $\beta\alpha$  IR) and actin intergenic region (Actin IR) processing signals. The total length of the ORF for *RAD51-6* is 1434 bp, and the size of the individual components (in bp), restriction sites used in the cloning and primer binding sites used to confirm the generation of the mutants are indicated.

### 3.5.2 Generating the *RAD51-6* mutant cell lines

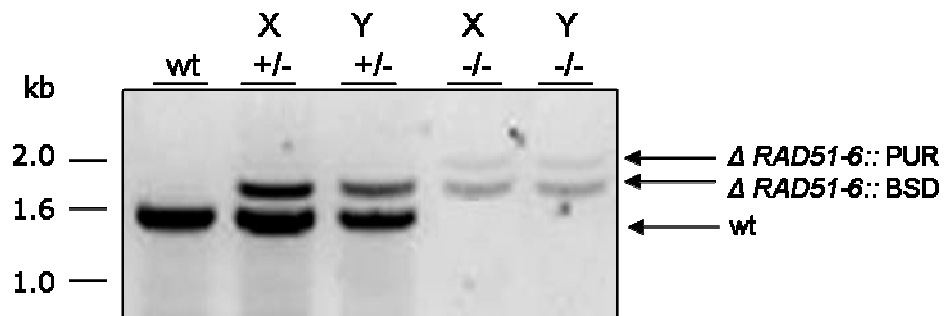
The  $\Delta RAD51-6::BSD$  and  $\Delta RAD51-6::PUR$  constructs were linearised using the restriction enzymes *Xba*I and *Xho*I. The DNA was phenol/chloroform extracted and ethanol precipitated. Two transformations using constructs  $\Delta RAD51-6::BSD$  were carried out to generate two independent *RAD51-6* heterozygous mutants, named X<sup>+/–</sup> and Y<sup>+/–</sup>. A total of  $5 \times 10^7$  *T. brucei* bloodstream stage 3174.2 cells were transformed by electroporation with approximately 5 µg of the digested  $\Delta RAD51-6::BSD$  construct. The cells were incubated at 37 °C for 24 hours in non-selective HMI-9. After the recovery period,  $5 \times 10^6$  cells were diluted in 36 ml of selective HMI-9 and plated over a 24 well plate at 1.5 ml per well, i.e. approximately  $2.08 \times 10^5$  cells per well. A blasticidin concentration of  $2.5 \mu\text{g} \cdot \text{ml}^{-1}$  was used for selection of the heterozygous mutant cell lines. Integration of the constructs into the expected gene locus was confirmed by PCR using a forward primer upstream of the ORF (Upstream of *RECA2* ORF For) and the reverse primer of the RHF (*RECA2* EndORF Rev); this data is not shown. Subsequently, to disrupt the second *RAD51-6* allele, a second round of transformation was carried out using constructs  $\Delta RAD51-6::PUR$  with the two independent cell lines, X<sup>+/–</sup> and Y<sup>+/–</sup>, to generate the homozygous cell lines, X<sup>–/–</sup> and Y<sup>–/–</sup>. The resistance marker for the second round of transformants was puromycin at a concentration of  $0.5 \mu\text{g} \cdot \text{ml}^{-1}$ . The X<sup>–/–</sup> and Y<sup>–/–</sup> cell lines were then confirmed by PCR-amplification, Southern blot, and reverse transcriptase-PCR analysis (see Section 3.6).

### 3.5.3 Confirmation of *RAD51-6* mutant cell lines

#### 3.5.3.1 Confirmation of *RAD51-6* mutants by PCR

As for *RAD51-4* (see Section 3.2), PCR-amplification of the entire *RAD51-6* ORF, using forward primers that bind upstream of the LHF (Upstream of *RECA2* ORF For; see Material and Methods, Section 2.4.1, Table 2-1) and the reverse primer of the RHF (*RECA2* EndORF Rev), was used to analyse the putative *RAD51-6* heterozygous (+/–) and homozygous mutants (–/–) *T. brucei* strain 3174.2. Figure 3-22 shows the primer binding sites and Figure 3-23 show the PCR products obtained from this PCR amplification, using *Taq* polymerase and genomic DNA as template from the wildtype (wt) 3174.2, the two *RAD51-6* +/– and two *RAD51-6* –/– cell lines. One product of the size expected for the intact *RAD51-6* ORF was amplified (1435 bp) from wildtype genomic DNA. From each of the *RAD51-6* +/– mutants, two products were amplified: one corresponded to the wildtype allele and the other to the size expected of the *BSD*-disrupted allele following integration of the antibiotic resistance cassettes into the ORF. The  $\Delta RAD51-6::BSD$  construct is approximately 350 bp larger than the deleted fragment of the *RAD51-6* ORF and the

resulting band is 1785 bp in length. From the two independent *rad51-6*  $-/-$  mutants, no wildtype PCR product was amplified and the two larger products are consistent with integration of both antibiotic cassettes, disrupting both alleles. The  $\Delta RAD51-6::PUR$  construct is 550 bp larger than the deleted fragment of the *RAD51-6* ORF and the resulting band is 1985 bp in length.



**Figure 3-23: Analysis of the *RAD51-6* gene disruption by PCR-amplification.**

An agarose gel separation of the PCR products generated by PCR amplification of the *RAD51-6* locus using the forward primer: Upstream of *RECA* ORF For and reverse primer: *RECA2* EndORF Rev is shown. PCR reaction were performed on genomic DNA from wildtype (wt) 3174.2 *T. brucei* cells, for two independent putative *RAD51-6* heterozygous mutants (X $+/$ - and Y $+/$ -) and two putative *rad51-6* homozygous mutants (X $-$ /- and Y $-$ /-). The DNA fragments of the size expected from the wt genes or from alleles disrupted by integration of the *BSD* resistance cassette ( $\Delta RAD51-6::BSD$ ) or the *PUR* resistance cassette ( $\Delta RAD51-6::PUR$ ) are indicated, as are the size markers.

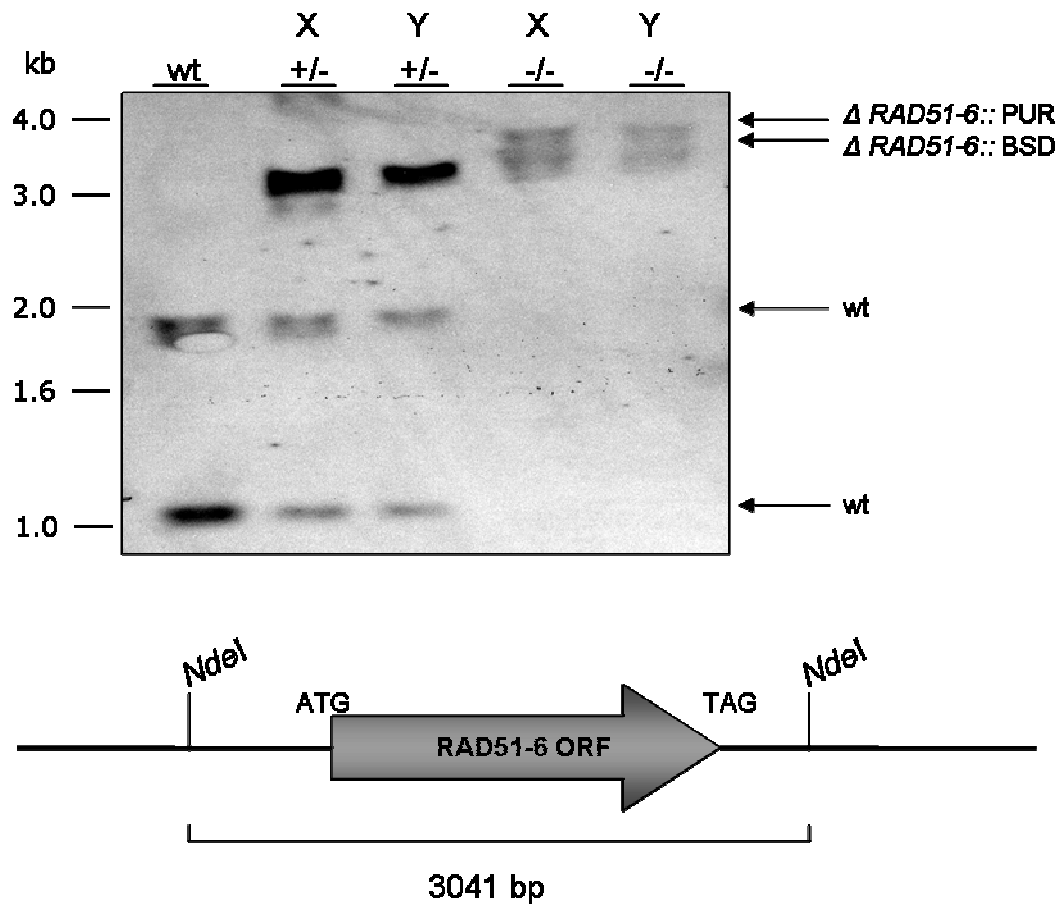
### 3.5.3.2 Confirmation of *RAD51-6* mutants by Southern blot analysis

Southern blot analysis was next carried out to confirm the PCR prediction of the generation of *RAD51-6* mutant cell lines (Materials and Methods, Section 2.5). The genomic DNA from each of the cell lines was extracted and digested with *NdeI*. This was then separated on a 0.8 % agarose gel and transferred to a nylon membrane, which was probed with the LHF of the *RAD51-6* ORF. The probe was made by PCR-amplifying the LHF using the primers *RECA2* StartORF For and *RECA2* StartORF Rev. The resulting PCR-product was then gel extracted, and labelled using the *Gene Images* CDP-*Star* Detection Module from Amersham.

Figure 3-24A shows the results of the Southern blot analysis. In wildtype cells, two products of ~1100 bp and ~1900 bp were obtained. It was predicted by sequence analysis that digestion with *NdeI* should generate one band of 3041 bp, (see Figure 3-24B). It is therefore possible that an allelic difference between the sequence TREU927 genomic DNA and Lister 427 genomic DNA used here resulted in the presence of an *NdeI* site. In the putative heterozygous mutants three hybridizing bands were obtained: two corresponded to the fragment observed in the wildtype cells (~1.1 kbp and ~1.9 kbp) and the other was of a



size consistent with the *BSD*-disrupted allele following integration of the *RAD51-6::BSD* construct into the *RAD51-6* ORF (3291 bp). In the putative homozygous mutants, both bands seen in the wildtype genomic DNA were absent and replaced with two larger fragments that are consistent in size with integration of both antibiotic cassettes, disrupting both *RAD51-6* alleles (*RAD51-6::BSD*: 3291 bp and *RAD51-6::PUR*: 3591 bp).



**Figure 3-24: Analysis of *RAD51-6* gene disruption by Southern blot analysis.**

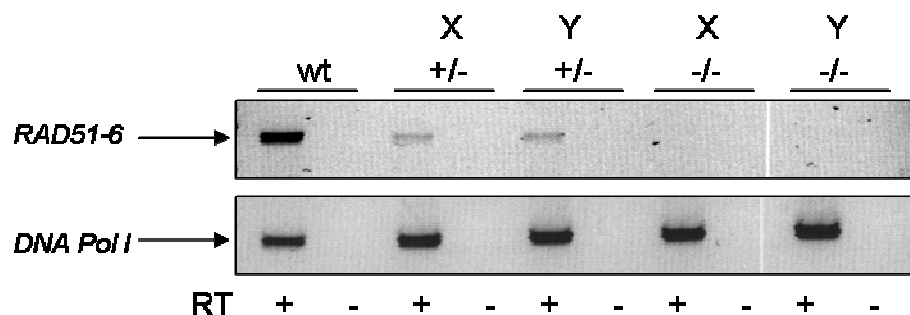
**Figure 3-24A:** Southern blot showing *NdeI*-digested genomic DNA and probed with the LHF of the *RAD51-6* ORF, see Figure 3-22. The extracted genomic DNA from wildtype (wt) 3174.2 *T. brucei* cells, from the two independent putative heterozygous mutants (X+/- and Y+/-) and from the two independent putative homozygous mutants (X-/- and Y-/-) was digested with *NdeI*, and separated on a 0.8 % agarose gel. The DNA was then transferred to a nylon membrane and probed with the LHF of the *RAD51-6* ORF. The DNA fragments of the size expected from the wt genes or from alleles disrupted by integration of the *BSD* resistance cassette ( $\Delta RAD51-6::BSD$ ) or the *PUR* resistance cassette ( $\Delta RAD51-6::PUR$ ) are indicated, as are the size markers. **Figure 3-24B:** shows the predicted locations of *NdeI* restriction enzyme sites around the *RAD51-6* ORF of the *T. brucei* strain TRUE927, predicted by sequence analysis (GeneDB).

### 3.5.3.3 Confirmation of *RAD51-6* mutants by RT-PCR

To confirm that mRNA of *RAD51-6* was not transcribed in the putative *rad51-6* -/- mutants, RT-PCR amplification was carried. Briefly, the total RNA was extracted from wt, *RAD51-6* +/-, and *rad51-6* -/- cells and converted into cDNA using random primers and reverse transcriptase (RT) enzyme (see Materials and Methods, Section 2.3). This cDNA was then used as template in the PCR reactions that amplified the entire *RAD51-6* ORF,

using a forward primer that binds to the LHF (*RecA2* StartORF For) and reverse primers that bind to the RHF (*RecA2* EndORF Rev).

Figure 3-25 shows the PCR fragments amplified from the cDNA. From wildtype and *RAD51-6* +/- cell lines, one product of the size expected for the *RAD51-6* ORF was amplified from the cDNA (1435 bp). From the *rad51-6* -/- cells, no *RAD51-6* mRNA was detected. DNA polymerase I RT-PCR positive controls confirmed that cDNA was present in the -/- cell samples, while RT- control confirmed that the PCR product generated from the cDNA and not from the contaminated genomic DNA.



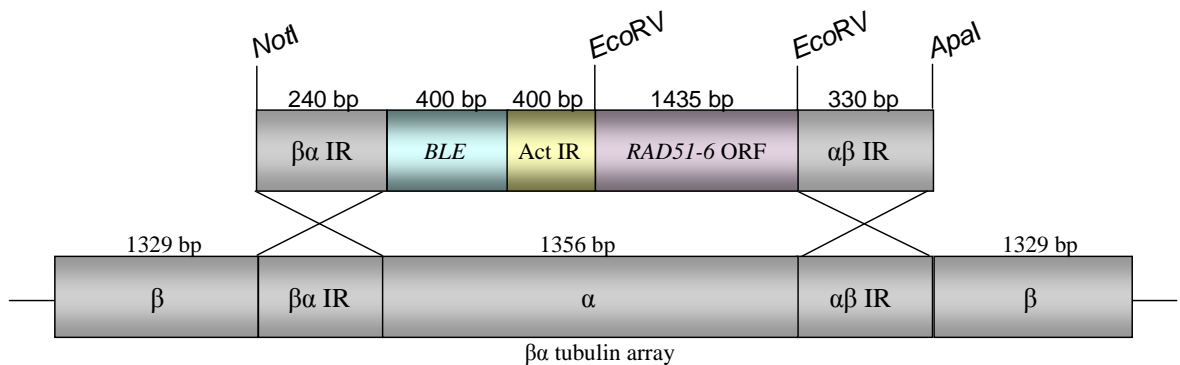
**Figure 3-25: Analysis of the expression of the *RAD51-6* mRNA in the mutant cell lines by RT-PCR amplification.**

Using the reverse transcriptase (RT) enzyme, cDNA was generated from mRNA extracted from the wildtype (wt) 3174.2 *T. brucei* cells, from the two independent putative *RAD51-6* heterozygous mutants (X+/- and Y+/-) and from the two independent putative *rad51-6* homozygous mutants (X-/- and Y-/-). This cDNA was used as template in the PCR reaction. PCR amplification was attempted using the forward primer (*RECA2* StartORF For) and reverse primer (*RECA2* EndORF Rev). The PCR products were then separated on a 1.0 % gel. As a positive control the DNA polymerase I ORF (*DNA Pol I*) was amplified from the cDNA of all the mutant cell lines. RT + and - indicates the presence of RT enzyme (+) and the absence of RT enzyme (-).

### 3.6 Generation of *RAD51-6* re-expressor cell lines

Having confirmed the generation of the *rad51-6* -/- mutants, it was necessary to re-express the *RAD51-6* gene product in these cell lines, in order to see if any phenotypes that arose were due to *RAD51-6* gene disruption (see Section 3.7) and not a secondary mutation. This was carried out using the same method to generate the *RAD51-4* re-expressor (see Section 3.3). The ORF of *RAD51-6* was PCR amplified using *RECA2* REX primers, (see Materials and Methods, Section 2.4.1, Table 2-1) and cloned using the restriction sites *EcoRV*, into pRM481 vector. As mentioned in Section 3.3 above, this cloning placed the *RAD51-6* ORF downstream of the phleomycin resistance cassette (BLE) and actin intergenic region and upstream of the  $\alpha\beta$  intergenic region in the pRM481 vector, ( $\beta\alpha$ -BLE-Act-*RAD51-6*- $\alpha\beta$ ). This construct was digested with *NotI* and *ApaI* and transformed into the *rad51-6* -/- mutant cell lines. Then the cells were allowed to recover for 24 hours in non-selective HMI-9 media and were plated on 2.5  $\mu\text{g}.\text{ml}^{-1}$  of phleomycin (Calbiochem). The flanking

target  $\beta\alpha$  and  $\alpha\beta$  regions allows integration of the *RAD51-6* and phleomycin resistance cassette by recombination into the tubulin locus with the removal of the  $\alpha$  tubulin gene and generation of the *RAD51-6* re-expressor cell lines (-/-/+; see Figure 3-26).

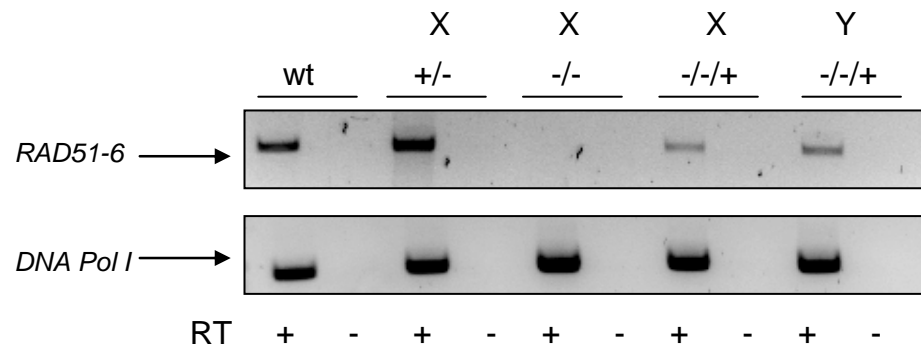


**Figure 3-26: Strategy for re-expression of the *RAD51-6* by recombination into the  $\beta\alpha$  tubulin array.**

The ORF of *RAD51-6* was cloned into the restriction site *EcoRV* of the pRM481 (top part of the diagram). *NotI* and *ApaI* digestion liberates the construct which was then transformed into the mutant cell lines and the integration of the *RAD51-6* ORF into the  $\beta\alpha$  tubulin array.

### 3.6.1 Confirmation of the *RAD51-6* re-expressor cell lines by RT-PCR

RT-PCR-amplification analysis, using the primers that bind to the *RAD51-6* LHF and RHF (*RECA2* StartORF For and *RECA2* EndORF Rev), was used to detect cDNA in the putative *RAD51-6* -/-/+ cell lines (see Material and Methods, Section 2.3). The PCR-amplification products are shown in Figure 3-27. One product of the size expected was amplified from the cDNA from the wildtype, *RAD51-6* X+/- and *RAD51-6* re-expressor (X-/-/+ and Y-/-/+) cell lines. From the *rad51-6* X-/- cDNA, no product was detected. DNA polymerase I was amplified from the cDNA of all the cell lines, confirming that the extracted RNA was converted to cDNA. The RT negative control had no RT enzyme (-) in the reaction and confirmed that the PCR products generated arose from the cDNA and not from contaminating genomic DNA.



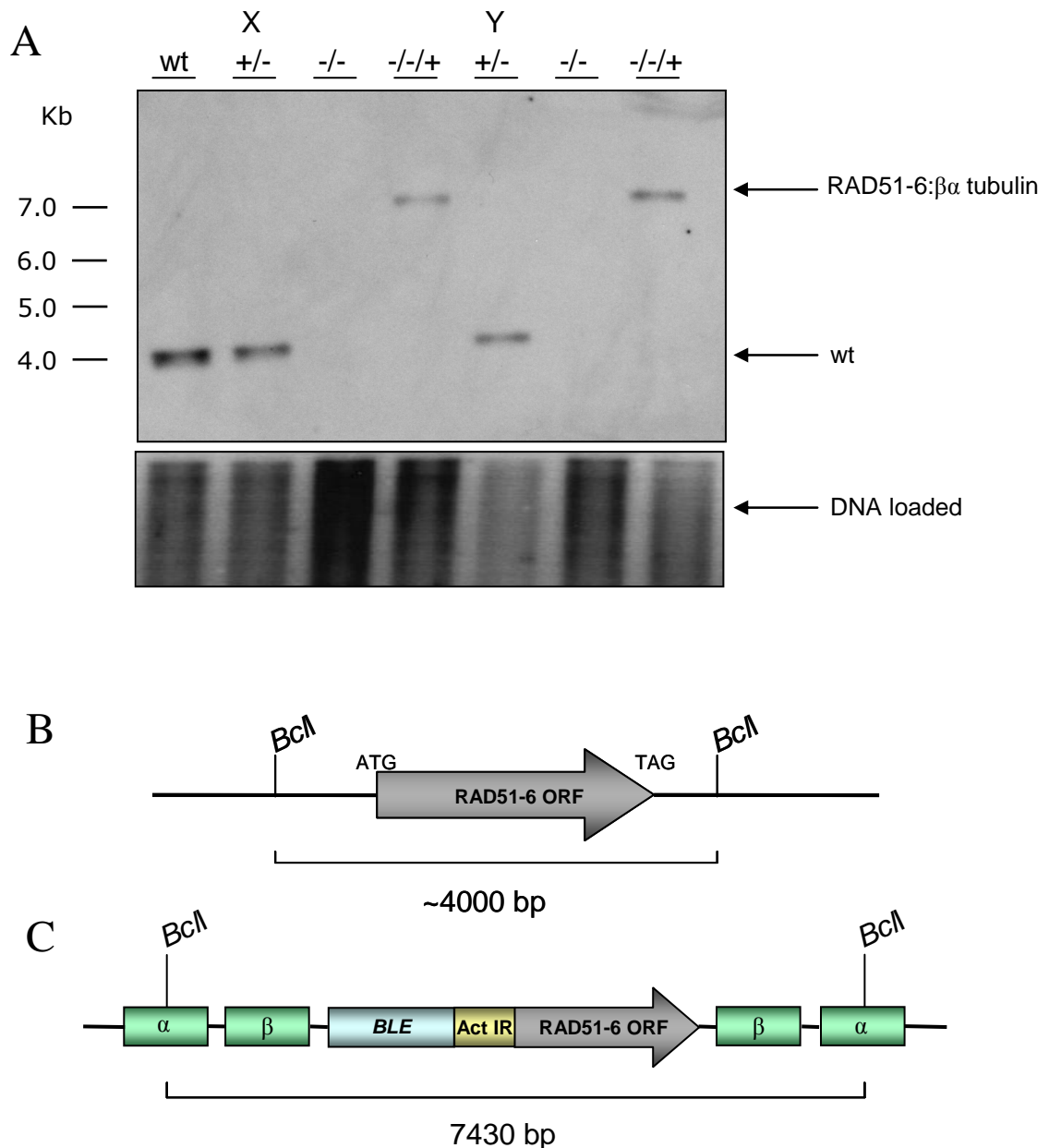
**Figure 3-27: Analysis of the expression of the *RAD51-6* mRNA in the *RAD51-6* re-expressor cell lines by RT-PCR amplification.**

Using the reverse transcriptase (RT) enzyme, cDNA was generated from mRNA extracted from the wildtype (wt) 3174.2 *T. brucei* cells, from the one putative *RAD51-6* heterozygous mutants (X+/-), from one putative *rad51-6* homozygous mutants (X-/-) and from the two independently generated *RAD51-6* re-expressor cell lines. This cDNA was used as template in the PCR reaction. PCR amplification was attempted using the forward primer (*RECA2StartORF* For) and reverse primer (*RECA2EndORF* Rev). The PCR products were then separated on a 1.0 % gel. As a positive control the DNA polymerase I ORF (*DNA Pol I*) was amplified from the cDNA of all the mutant cell lines. RT + and - indicates the presence of RT enzyme (+) and the absence of RT enzyme (-).

### 3.6.2 Confirmation of the *RAD51-6* re-expressors by Southern analysis

Southern analysis was carried out as a further test, to confirm the generation of *RAD51-6* -/-/+ cell lines (Materials and Methods, Section 2.5). The genomic DNA from the wildtype, *RAD51-6* +/-, *rad51-6* -/- and *RAD51-6* -/-/+ cell lines was extracted and digested with *Bcl*I. This was then separated on a 0.8 % agarose gel and transferred on to a nylon membrane. The membrane was probed with labelled DNA derived from the middle of the *RAD51-6* ORF, which was made by PCR-amplification using the primers PTP-*RAD51-6* For and *RAD51-6* middle Rev (see Materials & Methods, Section 2.4.1, Table 2-1), this product was then gel extracted, and labelled using the *Gene Images* CDP-*Star* Detection Module (Amersham).

Figure 3-28A shows the results of this Southern blot analysis. In wildtype and *RAD51-6* +/- cells, one fragment of approximately 4000 bp was observed, consistent with the prediction by sequence analysis, as the endogenous *RAD51-6* has two *Bcl*I sites either side of it (see Figure 3-28B). In the *rad51-6* -/- mutants, there were no bands visible, although DNA was loaded into the gel, as shown by the UV picture of the gel before transfer to the membrane. In the re-expressor cell lines, the *RAD51-6* ORF was integrated into the  $\beta\alpha$  tubulin array, resulting in a different restriction pattern: *Bcl*I digests the array in the  $\alpha$  tubulin gene, and resulted in a 7430 bp fragment which is larger than the endogenous *RAD51-6* band (see Figure 3-28C),



**Figure 3-28: Analysis of the generation of the *RAD51-6* re-expressor cell lines by Southern blot.**

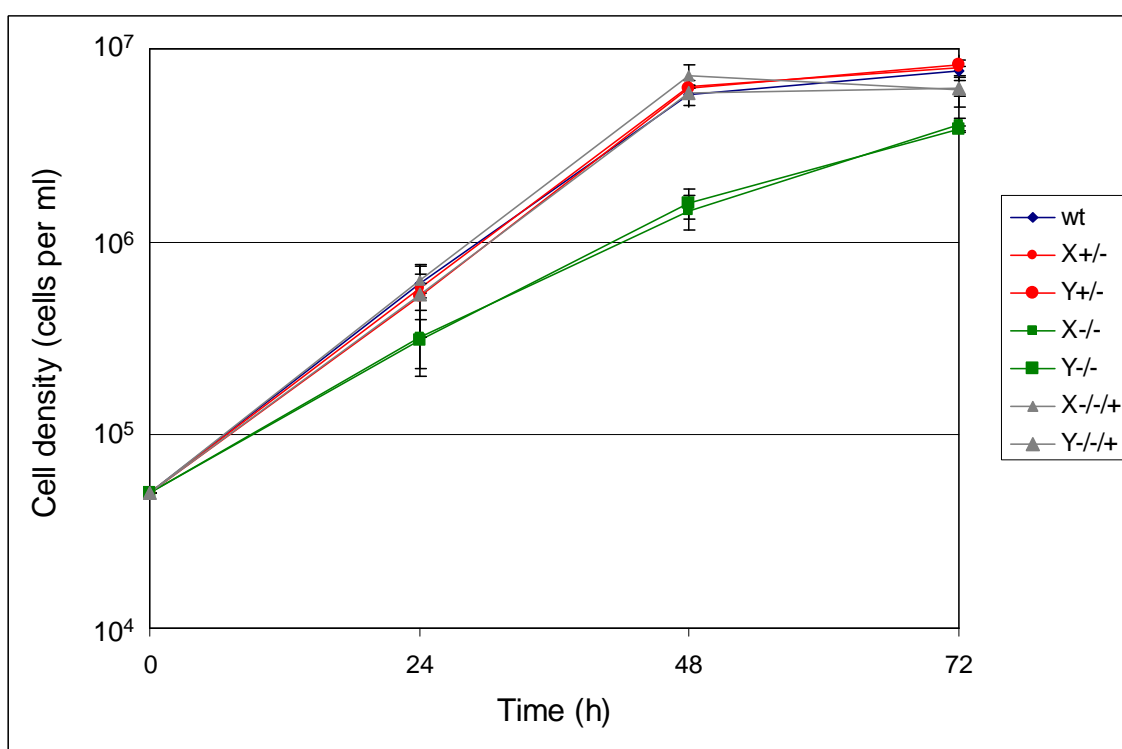
**Figure 3-28A:** Southern blot showing *BclI*-digested genomic DNA and probed with the middle of the *RAD51-6* ORF. The extracted genomic DNA from wildtype (wt) 3174.2 *T. brucei* cells, from the two independent putative heterozygous mutants (X<sup>+/-</sup> and Y<sup>+/-</sup>), from the two independent putative homozygous mutants (X<sup>-/-</sup> and Y<sup>-/-</sup>) and from the two independent putative *RAD51-6* re-expressor (X<sup>-/-/+</sup> and Y<sup>-/-/+</sup>) was digested with *BclI*, and separated on a 0.8 % agarose gel. The DNA was then transferred to a nylon membrane and probed with the middle of the *RAD51-6* ORF. The DNA fragments of the size expected from the wt gene, from alleles disrupted by integration of the *BSD* resistance cassette ( $\Delta$ *RAD51-6::BSD*) and integration of re-expressor *RAD51-6* into the tubulin array (*RAD51-6::β $\alpha$* ) are indicated, as are the size markers. **Figure 3-28B:** shows the locations of *BclI* restriction enzyme sites up and down stream of the *RAD51-6* ORF. **Figure 3-28C:** shows the locations of *BclI* site in the  $\alpha$ -tubulin array predicted by sequence analysis.

## 3.7 Phenotypic analysis of *RAD51-6* mutants and re-expressor cell lines

Phenotypic analysis of the *RAD51-6* mutant cell lines was carried out in a similar manner to that of the *RAD51-4* mutant cell lines. To avoid repetition, see the corresponding Sections above (Section 3.4) and associated Materials and Methods for description of experimental details (Chapter 2) and Appendices for statistical analysis (Appendix 3).

### 3.7.1 Analysis of *in vitro* growth rate of *RAD51-6* mutant cell lines

To determine if mutation of *RAD51-6* had any effect on the growth of *T. brucei*, population doubling time were measured, as described in Section 3.4.1. Briefly, to examine growth rates, each cell line was inoculated at  $5 \times 10^4$  cells per ml into HMI-9 medium and cell densities were counted at fixed time points. The results of this are plotted on a semi-logarithmic graph in Figure 3-29 and the population doubling times shown in Table 3-10.



**Figure 3-29: Analysis of growth of *RAD51-6* mutant cell lines *in vitro*.**

The semi-logarithmic of cell densities were plotted against fixed time points, (24 h, 48 h, and 72 h) is shown. Wildtype 3174.2 strain (wt) is compared with two independent *RAD51-6* heterozygous mutants (X+/-, Y+/-), *rad51-6* homozygous mutants (X-/-, Y-/-) and *RAD51-6* re-expressor cell lines (X-/-+, Y-/-+). The graph shows the means of triplicate data, with the 95 % confidence intervals indicated.

**Table 3-10: *in vitro* doubling times for *RAD51-6* mutant cell lines.**

The population doubling time for the two independently generated *RAD51-6* heterozygote (X+/-, Y+/-), *rad51-6* homozygous mutants (X-/-, Y-/-) and *RAD51-6* re-expressor cell lines (X-/-/+, Y-/-/+) are compared with the doubling time for wildtype strain 3174.2 (wt). The data are the means of the triplicate experiments (in hours) with the 95 % confidence intervals indicated (in brackets).

wt	X +/-	Y +/-	X -/-	Y +/-	X -/-/+	Y -/-/+
6.93 (+/-0.05)	6.90 (+/-0.03)	6.95 (+/-0.18)	9.98 (+/-0.63)	9.72 (+/-0.50)	6.73 (+/-0.18)	7.01 (+/-0.20)

These data show that *rad51-6* -/- mutants had a reduced growth rate, with each independent mutant X-/- and Y-/-, displaying a decrease in growth of 25.3 % and 27.5 % compared to the wildtype at the 48 h time point (Figure 3-29). This was confirmed by increased population doubling times compared with the wildtype and *RAD51-6* +/- mutants. *rad51-6* X-/- have a population doubling time of 9.98 h (+/- 0.63) and *rad51-6* Y-/- have a population doubling time of 9.72 h (+/- 0.50), compared with the wildtype's 6.93 h (+/- 0.05); see Table 3-10. These results are significantly different as evaluated by the Student's *t*-tests for the population doubling time (Table 3-11; all other pair-wise statistical comparisons are shown in Appendix 3, Table 8). Re-expression of the *RAD51-6* in each of the homozygous mutant cell lines enabled the cells to recover from the growth deficiency. This confirmed that lack of *RAD51-6* was responsible for the slow growth of the cell lines. Interestingly, the *rad51-6* -/- mutant was also slower growing than the *rad51-4* -/- mutants; this is discussed in Chapter 4 (Section 4.2).

**Table 3-11: Statistical analysis of the *in vitro* population doubling times for *RAD51-6* mutants.**

Unpaired, two-tailed Student's *t*-tests were carried out to compare the mean of the triplicate doubling time results for the two independently generated *RAD51-6* heterozygote (X+/-, Y+/-) and *rad51-6* homozygous mutants (X-/-, Y-/-) and *RAD51-6* re-expressor cell lines (X-/-/+, Y-/-/+) against the wildtype. Significant difference between the means is shown by a *p*-value of  $p \leq 0.05$ ; Not significantly different results are shown as "-". [NA is not applicable.]

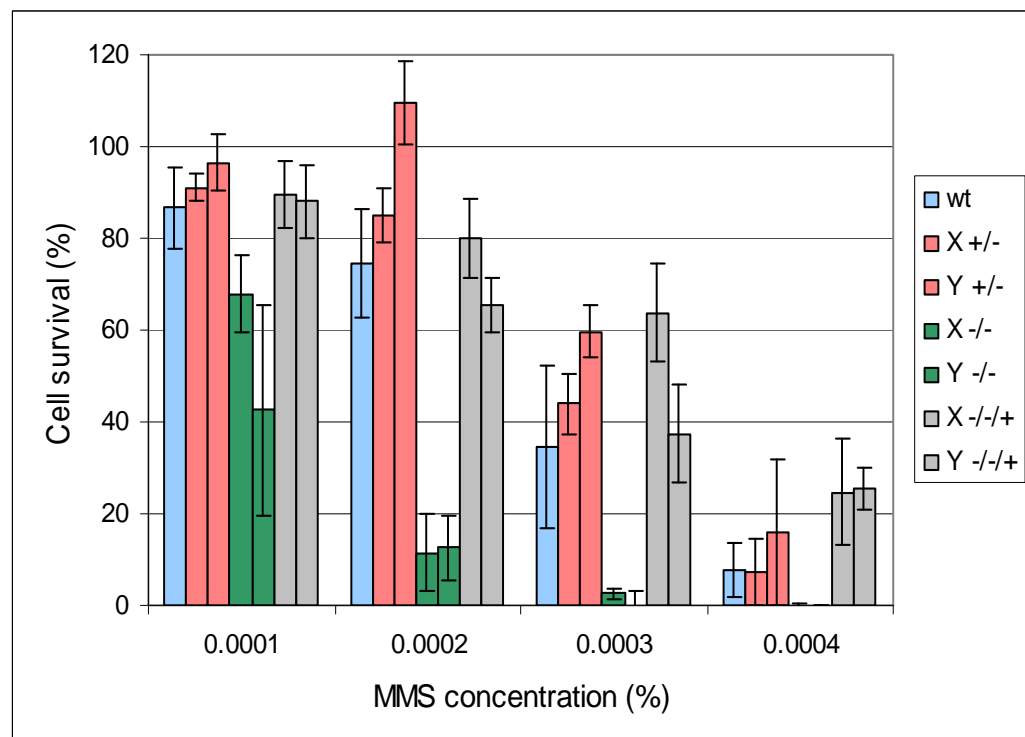
	wt	X +/-	Y +/-	X -/-	Y -/-	X -/-/+	Y -/-/+
wt	NA	-	-	0.0007	0.0004	-	-

### 3.7.2 Analysis of DNA damage sensitivity of *RAD51-6* mutant cell lines

To analyse if *RAD51-6* mutation lead to increased DNA-damage sensitivity, again two different assays were carried out: clonal survival in the presence of DNA-damaging agents, and measuring the metabolic capacity through absorbance of resazurin to determine the IC<sub>50</sub> values for each cell line exposed to the DNA damaging agents, MMS and phleomycin.

### 3.7.2.1 Clonal Survival of *RAD51-6* mutant cell lines

As for the *RAD51-4* mutant cell lines (see Section 3.4.2.1), the clonal survival assay involved diluting  $2 \times 10^6$  cells per ml to a concentration of 1 cell per well over a 96 well plate and determining the number of wells that yield growing populations in the presence of varying concentrations of MMS (%) and phleomycin ( $\mu\text{g} \cdot \text{ml}^{-1}$ ). Figure 3-30 shows the clonal survival of *rad51-6*  $-/-$  mutants compared with that of the wildtype, *RAD51-6*  $+/-$  and *RAD51-6*  $-/-/+$  cell lines when treated with 0.0001 %, 0.0002 %, 0.0003 %, and 0.0004 % of MMS. The data is represented as a percentage of growth when treated with MMS compared to growth its absence.



**Figure 3-30: Clonal survival of *RAD51-6* mutants *in vitro* when treated with methyl methanesulfonate (MMS).**

Cell survival is shown as a percentage of clonal growth when treated with MMS compared to growth in the absence of MMS. Each cell line wildtype 3174.2 (wt), two independent *RAD51-6* heterozygous mutants (X $+/-$ , Y $+/-$ ), *rad51-6* homozygous mutants (X $-/-$ , Y $-/-$ ) and re-expressor cell lines (X $-/-/+$ , Y $-/-/+$ ) was treated with increasing concentration of MMS (0.0001 %, 0.0002 %, 0.0003 %, 0.0004 %). The graph shows the means of triplicate data, with 95 % confidence intervals indicated by error bars.

With increasing concentrations of MMS, there was a clear decrease in survival of each *rad51-6*  $-/-$  mutant, X $-/-$  and Y $-/-$ , compared with all other cell lines. At a MMS concentration of 0.0002 %, for instance, wildtype survival was 74.4 % ( $\pm$  17.6) relative to the untreated cells, while the *rad51-6* X $-/-$  displayed 11.5 % ( $\pm$  1.1) survival and *rad51-6* Y $-/-$  displayed 12.5 % ( $\pm$  3.3) survival. The results of unpaired, two tailed Student's *t*-tests for clonal survival at this MMS concentration is shown in Table 3-12, confirming that



clonal survival of each *rad51-6*  $-/-$  mutant is significantly reduced compared with the wildtype cells.

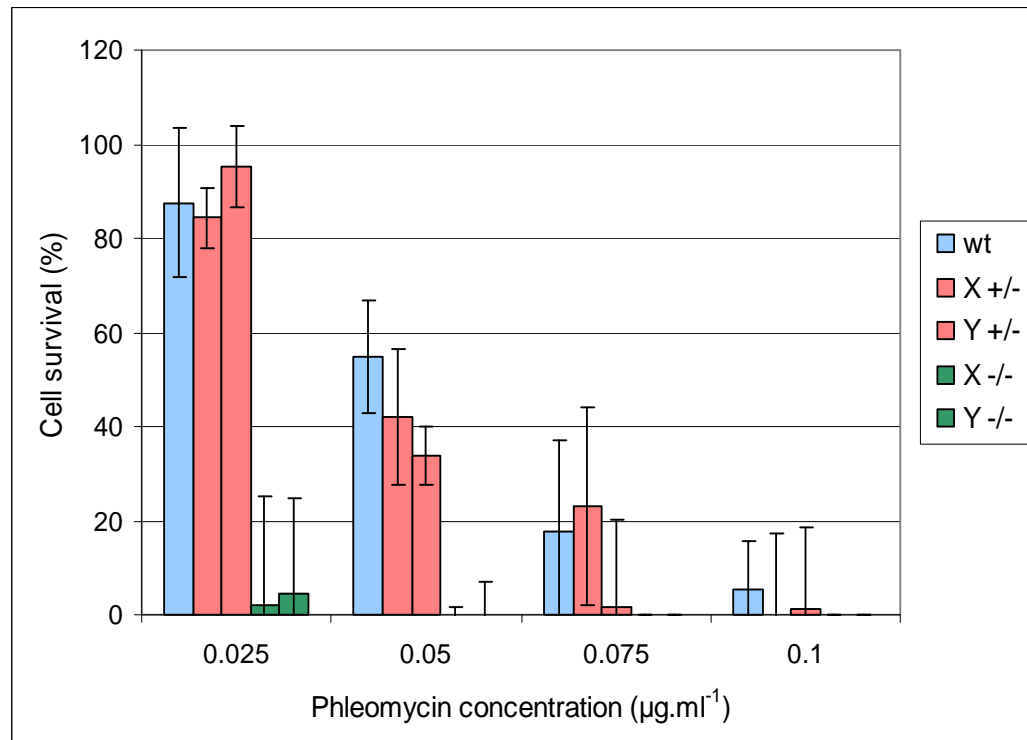
**Table 3-12: Statistical analysis of the clonal survival of *RAD51-6* mutants when treated with 0.0003 % methyl methanesulfonate (MMS).**

Unpaired two-tailed Student's *t*-tests were carried out to compare the mean of the triplicate clonal survival results for the cell lines against the wildtype. Significant difference between the results is shown by a *p*-value of  $p \leq 0.05$ , while not significantly different results are shown as “-”. [NA is not applicable] Each cell line is indicated as follows: wildtype strain 3174.2 (wt), two independent *RAD51-6* heterozygous mutants (X+/-, Y+/-), *rad51-6* homozygous mutants (X-/-, Y-/-) and *RAD51-6* re-expressor cell lines (X-/-/+, Y-/-/+).

	wt	X+/-	Y+/-	X-/-	Y-/-	X-/-/+	Y-/-/+
wt	NA	-	-	0.0043	0.0132	-	-

Student's *t*-tests were also carried out to compare the mean survival rate if all the cell lines against each other at all different concentrations of MMS, and these data are shown in Appendix 3, Table 9. These results show that with increasing concentration of MMS, there was a significant reduction of clonal survival of the *rad51-6*  $-/-$  mutants compared with wildtype and *RAD51-6*  $+/-$  cells. In addition, re-expression of the functional *RAD51-6* in each of homozygous mutant cell line caused survival to increase.

Clonal survival data generated following treatment with four concentrations of phleomycin  $0.025 \mu\text{g.ml}^{-1}$ ,  $0.05 \mu\text{g.ml}^{-1}$ ,  $0.075 \mu\text{g.ml}^{-1}$ , and  $0.1 \mu\text{g.ml}^{-1}$  are shown in Figure 3-31.



**Figure 3-31: Clonal survival of *RAD51-6* mutants *in vitro* when treated with phleomycin.**

Cell survival is shown as a percentage of clonal growth when treated with phleomycin compared to growth in the absence of drug. Each cell line, wildtype 3174.2 (wt), two independent *RAD51-6* heterozygous mutants (X+/-, Y+/-), and *rad51-6* homozygous mutants (X-/-, Y-/-) was treated with increasing concentration of phleomycin (0.025 µg.ml<sup>-1</sup>, 0.05 µg.ml<sup>-1</sup>, 0.075 µg.ml<sup>-1</sup>, 0.1 µg.ml<sup>-1</sup>). The graph shows the means of triplicate data, with 95 % confidence intervals indicated by error bars.

As for the MMS, increasing concentrations of phleomycin from 0.025 µg.ml<sup>-1</sup> to 0.075 µg.ml<sup>-1</sup>, caused a clear decrease in survival of each *rad51-6* cell line, X-/- and Y-/-, compared with the other cell lines. At a phleomycin concentration of 0.025 µg.ml<sup>-1</sup>, wildtype survival was 87.6 % (+/- 16.0) compared with 1.9 % (+/- 23.2) and 4.5 % (+/- 20.0) for *rad51-6* X-/- and Y-/- respectively. The statistical significance of this finding is shown in Table 3-13.

**Table 3-13: Statistical analysis of the clonal survival for *RAD51-6* mutants when treated with 0.025 µg.ml<sup>-1</sup> phleomycin.**

Unpaired, two-tailed Student's *t*-tests were carried out to compare the mean of the triplicate clonal survival results for the cell lines against the wildtype, a significant difference between the results is shown by a *p*-value of *p* ≤ 0.05, while a non-significantly different results are shown as “-“. [NA is not applicable.] Each cell line is indicated as follows: wildtype 3174.2 (wt), two independent *RAD51-6* heterozygous mutants (X+/-, Y+/-), and *rad51-6* homozygous mutants (X-/-, Y-/-).

	wt	X+/-	Y+/-	X-/-	Y-/-
wt	NA	-	-	0.0002	0.0003

Unpaired, two-tailed Student's *t*-tests to compare the survival rates of all cell lines relative to each other, at all the different concentrations of phleomycin are shown in Appendix 3,

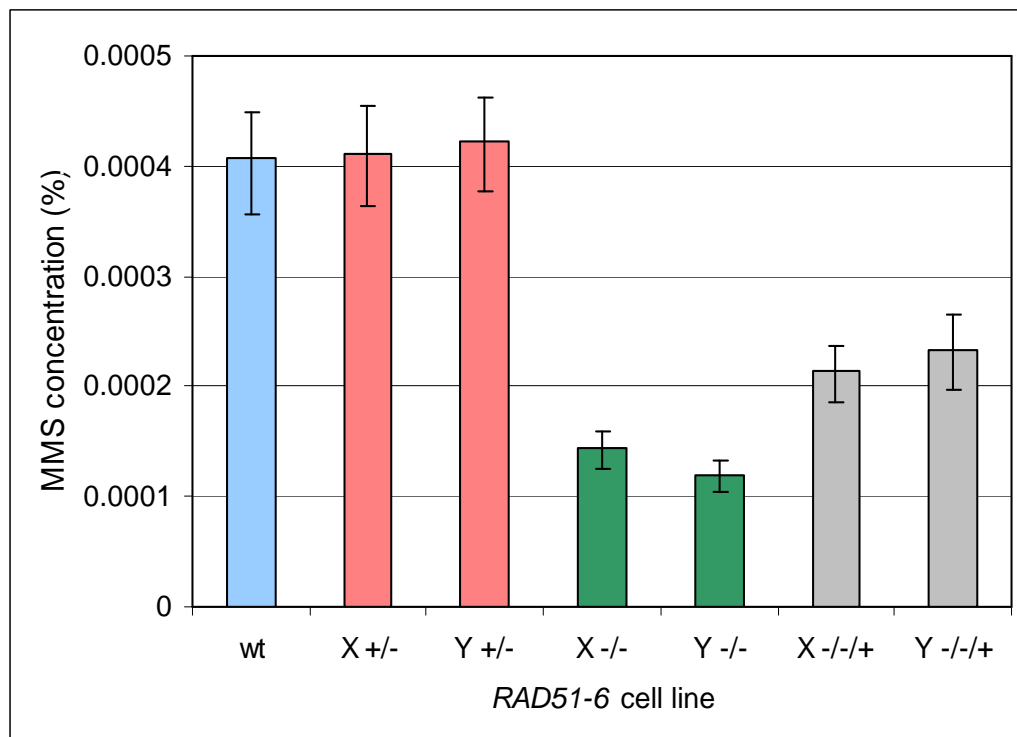
Table 10. At a concentration of  $0.05 \mu\text{g}.\text{ml}^{-1}$  of phleomycin, there was also a significant decrease of clonal survival of the *rad51-6*  $-/-$  mutant compared with wildtype and *RAD51-6*  $+/-$  mutants. However, at the higher concentrations of  $0.075 \mu\text{g}.\text{ml}^{-1}$  and  $0.1 \mu\text{g}.\text{ml}^{-1}$ , this analysis was not possible, as the *rad51-6*  $-/-$  mutants were killed. Re-expressor cell lines were not analysed, as functional add-back of the *RAD51-6* gene was selected by phleomycin resistance (Section 3.6).

These data show that the *rad51-6*  $-/-$  mutants, like *rad51-4*  $-/-$  mutants, are more sensitive to both DNA-damaging agents, MMS and phleomycin, than the wildtype and heterozygous cell lines.

### **3.7.2.2 *IC<sub>50</sub> values of MMS and Phleomycin for RAD51-6 mutant cell lines***

Determination of the  $\text{IC}_{50}$  values of MMS and phleomycin of the *RAD51-6* mutant cell lines was conducted by measuring conversion of resazurin to resorufin as described as for the *RAD51-4* mutant cell lines in Section 3.4.2.2. The same starting concentrations of cells, titration of MMS and phleomycin, and growth and measurement by spectrometry were performed (see Materials and Methods, Section 2.8.2.2).

Figure 3-32 shows the calculated mean  $\text{IC}_{50}$  values for MMS of wildtype, *RAD51-6*  $+/-$ , *rad51-6*  $-/-$  and *RAD51-6*  $-/-/+$  cell lines. The wildtype cells mean  $\text{IC}_{50}$  value was  $0.000406\%$  ( $\pm 0.000046$ ), which is 2.5 – 4.0 fold higher than that of the two *rad51-6*  $-/-$  mutants:  $0.000143\%$  ( $\pm 0.000017$ ) and  $0.000119\%$  ( $\pm 0.000015$ ) for X $-/-$  and Y $-/-$ , respectively. In contrast, each *RAD51-6*  $+/-$  mutant's  $\text{IC}_{50}$  value was equivalent to the wildtype cells.



**Figure 3-32: IC<sub>50</sub> values of MMS for RAD51-6 mutant cell lines**

Resazurin was used as an indicator of growth and IC<sub>50</sub> values were calculated for each cell line. Wildtype 3174.2 (wt), two independent *RAD51-6* heterozygous mutants (X+/-, Y+/-), two independent *rad51-6* homozygous mutants (X-/-, Y-/-) and two independent *RAD51-6* re-expressor cell lines (X-/-+, Y-/-+) were treated with serial dilutions of MMS. The graph shows the mean IC<sub>50</sub> values derivative from the triplicate data, with the 95 % confidence intervals indicated by the error bars.

The reduction of the IC<sub>50</sub> value of the *rad51-6* -/- mutants compared with the wildtype cells was significantly different, based on the unpaired, two-tailed Student's *t*-test comparisons of the data (Table 3-14). Student's *t*-tests of all other pair-wise comparisons are shown in Appendix 3, Table 11.

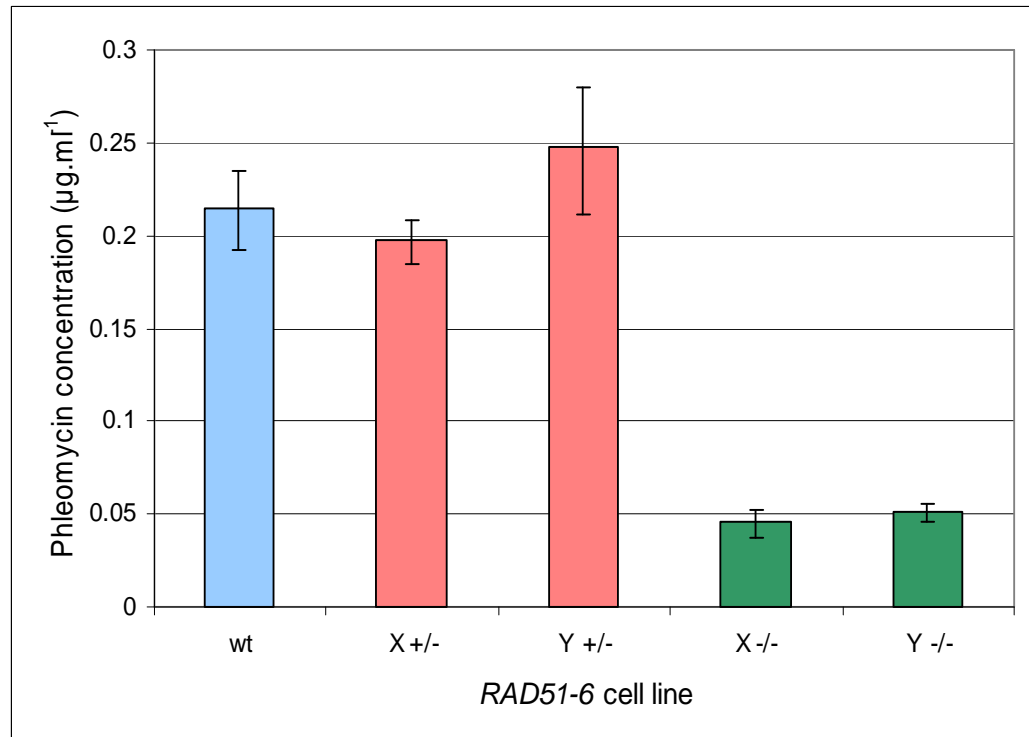
Perhaps surprisingly, re-expression of functional RAD51-6 in the *RAD51-6* -/-+ cell lines appeared not to revert the MMS sensitivity to equivalent levels of resistance as found in the wildtype or *RAD51-6* +/- cells. The reason for this is unclear, in particular as the same findings were not the case in the clonal survival assay, see Figure 3-30.

**Table 3-14: Statistical analysis of the IC<sub>50</sub> values for MMS of the RAD51-6 mutants.**

Student's *t*-tests were carried out to compare the mean of the triplicate IC<sub>50</sub> values for the cell lines against the wildtype. A significant difference between the results is shown by a *p*-value of  $p \leq 0.05$  and not significantly different results are shown as "-". [NA is not applicable.] Each cell line is indicated as follows, wildtype strain 3174.2 (wt), two independent *RAD51-6* heterozygous mutants (X+/-, Y+/-), two independent *rad51-6* homozygous mutants (X-/-, Y-/-) and two independent *RAD51-6* re-expressors (X-/-+, Y-/-+).

	wt	X +/-	Y +/-	X -/-	Y -/-	X -/- +	Y -/- +
wt	NA	-	-	0.0048	0.0010	-	-

Figure 3-33 shows the mean  $IC_{50}$  values calculated for phleomycin. Here, the wildtype cells display a mean  $IC_{50}$  value of  $0.214 \mu\text{g.ml}^{-1}$  ( $\pm 0.021$ ) and was approximately four fold higher than that of *rad51-6*  $-/-$  mutants whose mean  $IC_{50}$  values were  $0.046 \mu\text{g.ml}^{-1}$  ( $\pm 0.007$ ) and  $0.051 \mu\text{g.ml}^{-1}$  ( $\pm 0.004$ ) for X $-/-$  and Y $-/-$  respectively.



**Figure 3-33:  $IC_{50}$  values of phleomycin for RAD51-6 cell lines**

Resazurin blue was used as an indicator of growth and  $IC_{50}$  values were calculated for each cell line. Wildtype 3174.2 (wt), two independent RAD51-6 heterozygous mutants (X+/-, Y+/-), two independent *rad51-6* homozygous mutants (X-/-, Y-/-) and two independent RAD51-6 re-expressor cell lines (X-/-+, Y-/-+) were treated with serial dilutions of phleomycin. The graph shows the means of the triplicate data, with the 95 % confidence intervals indicated by the error bars.

Student's *t*-tests (Table 3-15) confirm that the reduction of the  $IC_{50}$  value for phleomycin of the wildtype cells compared with the *rad51-6*  $-/-$  mutants was significant. In contrast, no difference in phleomycin sensitivity was found for the RAD51-6 +/- cells relative to the wildtype cells. The complete statistical analysis of all pair-wise comparisons can be found in Appendix 3, Table 12.

**Table 3-15: Statistical analysis of the  $IC_{50}$  values for phleomycin of the RAD51-6 mutants.**

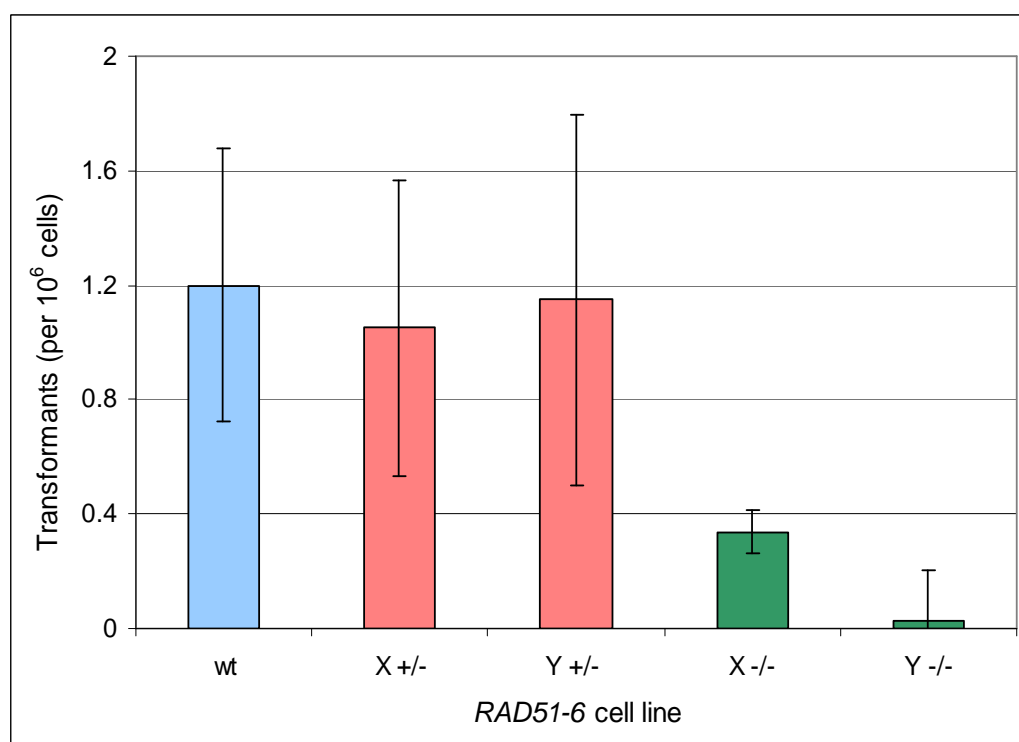
Student's *t*-tests were carried out to compare the mean of the triplicate  $IC_{50}$  values for the cell lines against the wildtype. A significant difference between the results is shown by a *p*-value of  $p \leq 0.05$  and not significantly different results are shown as "-". [NA is not applicable.] Each cell line is described: wildtype strain 3174.2 (wt), two independent RAD51-6 heterozygous mutants (X+/-, Y+/-), and two independent *rad51-6* homozygous mutants (X-/-, Y-/-).

	wt	X+/-	Y+/-	X-/-	Y-/-
wt	NA	-	-	0.0060	0.0108

These results largely confirm the findings of the clonal survival assays and show that survival in the presence of two DNA-damaging agents is reduced for *rad51-6*  $-/-$  mutants relative to wildtype cells and *RAD51*  $+/-$  mutants. Most likely, the lack of functional RAD51-6 in the cell causes an impairment in the parasite's ability to repair the DNA damage.

### 3.7.3 Analysis of recombination efficiency of *RAD51-6* mutant cell lines

To assay for recombination efficiency, a transformation assay was used, following the procedure described for *RAD51-4* mutants, see Section 3.4.3. The results of these experiments are shown in Figure 3-34.



**Figure 3-34: Transformation assay for the *RAD51-6* mutant cell lines.**

To assay for recombination rate, a transformation efficiency assay was used. Each cell line is indicated as follows: Wildtype strain 3174.2 (wt), two independent *RAD51-6* heterozygous mutants (X $+/-$ , Y $+/-$ ), and two independent *rad51-6* homozygous mutants (X $-/-$ , Y $-/-$ ). The values plotted are mean numbers of phleomycin resistant transformants per  $10^6$  cells transformed. These experiments were done in triplicate with the same linearised DNA, and error bars of the 95 % confidence intervals are indicated.

These data indicate that the *rad51-6*  $-/-$  mutants had a defect in transformation efficiency compared with the wildtype and *RAD51-6*  $+/-$  mutant cell lines. The transformation efficiency for the wildtype cells was  $1.2 \times 10^{-6}$  ( $\pm 3.4 \times 10^{-7}$ ) whereas the transformation efficiencies of the *rad51-6* X $-/-$  and Y $-/-$  were  $3.3 \times 10^{-7}$  ( $\pm 7.3 \times 10^{-8}$ ) and  $2.4 \times 10^{-7}$  ( $\pm 1.8 \times 10^{-7}$ ) respectively. These results were confirmed to be significantly different by

unpaired, two-tailed Student's *t*-tests (Table 3-16, and Appendix 3, Table 13). In contrast, the *RAD51-6* +/- mutants were as efficient as the wildtype cells in transformation.

**Table 3-16: Statistical analysis of the transformation efficiency of *RAD51-6* mutant cell lines.**

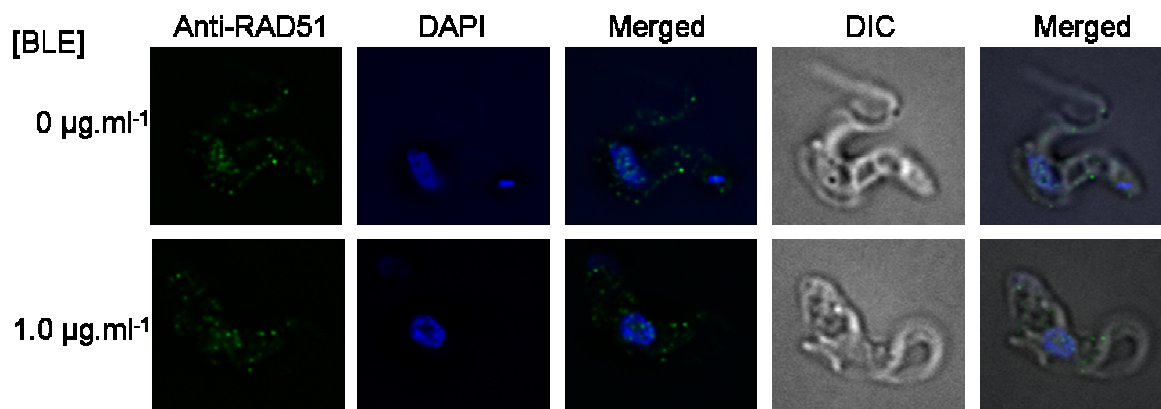
Unpaired, two-tailed Student's *t*-tests were carried out to compare the mean of the triplicate transformation efficiency results for the wildtype cells relative to the other cell lines. A significant difference between the results is shown by a *p*-value of  $p \leq 0.05$  and not significantly different results are shown as “-”. [NA is not applicable.] Each cell line is described: wildtype strain 3174.2 (wt), two independent *RAD51-6* heterozygous mutants (X+/-, Y+/-), and two independent *rad51-6* homozygous mutants (X-/-, Y-/-).

	wt	X+/-	Y+/-	X-/-	Y-/-
wt	NA	-	-	0.0220	0.0148

This analysis has shown that mutation of *RAD51-6* causes a decrease in the ability of *T. brucei* to undergo homologous recombination, a finding that is similar to that obtained for the *rad51-4* -/- mutants, see Section 3.4.3. It is also comparable with the finding for *rad51*, *rad51-3*, and *rad51-5* homozygous mutant cell lines (McCulloch & Barry, 1999; Proudfoot & McCulloch, 2005a). As a result, therefore, it appears that all four *T. brucei* *RAD51* paralogues act in HR.

### 3.7.4 Analysis of *RAD51* foci formation in *RAD51-6* mutant cell lines

In Section 3.4.4, it was demonstrated that DNA damage by phleomycin treatment of *T. brucei* causes a re-localisation of *RAD51* into distinct foci, and that *RAD51-3*, *RAD51-5* and *BRCA2* all contribute to the formation or stabilisation of *RAD51* foci (see Chapter 4, Section 4.5 of this thesis, Hartley & McCulloch, 2008, Proudfoot & McCulloch, 2005b). To examine if *RAD51-6* also acts in the *RAD51* relocalisation, immuno-fluorescence was carried out on fixed phleomycin treated and control non-treated cells using rabbit anti-*RAD51* antiserum, as described in Section 3.4.4 and Materials & Methods, Section 2.9. In this case a more limited range of phleomycin concentrations were examined: growth of *RAD51-6* +/-, and *rad51-6* -/- for 18 hours in either 0.25  $\mu\text{g}.\text{ml}^{-1}$  or 1.0  $\mu\text{g}.\text{ml}^{-1}$  phleomycin. Wildtype data are those generated previously for comparison with the *RAD51-4* cell lines (Section 3.4.4). The *RAD51-6* +/- cell lines displayed a similar pattern to wildtype, with 70-75 % of cells forming discernable *RAD51* foci at 1.0  $\mu\text{g}.\text{ml}^{-1}$ , compared with 78 % of the wildtype cells.



**Figure 3-35: RAD51 sub-nuclear foci in *rad51-6*  $-/-$  mutant.**

The first row of pictures shows an example of a non-treated *rad51-6*  $-/-$  cell: no distinct RAD51 foci are detected. The second row is an example of a *rad51-6*  $-/-$  cell treated with  $1.0 \mu\text{g.ml}^{-1}$  phleomycin for 18 hours. RAD51 was imaged by indirect immuno-fluorescence (FITC) (first column), the DNA was stained with DAPI (second column), FITC and DAPI were merged (third column), the cells are visualised by DIC (fourth column) and all the images were merged (fifth column).

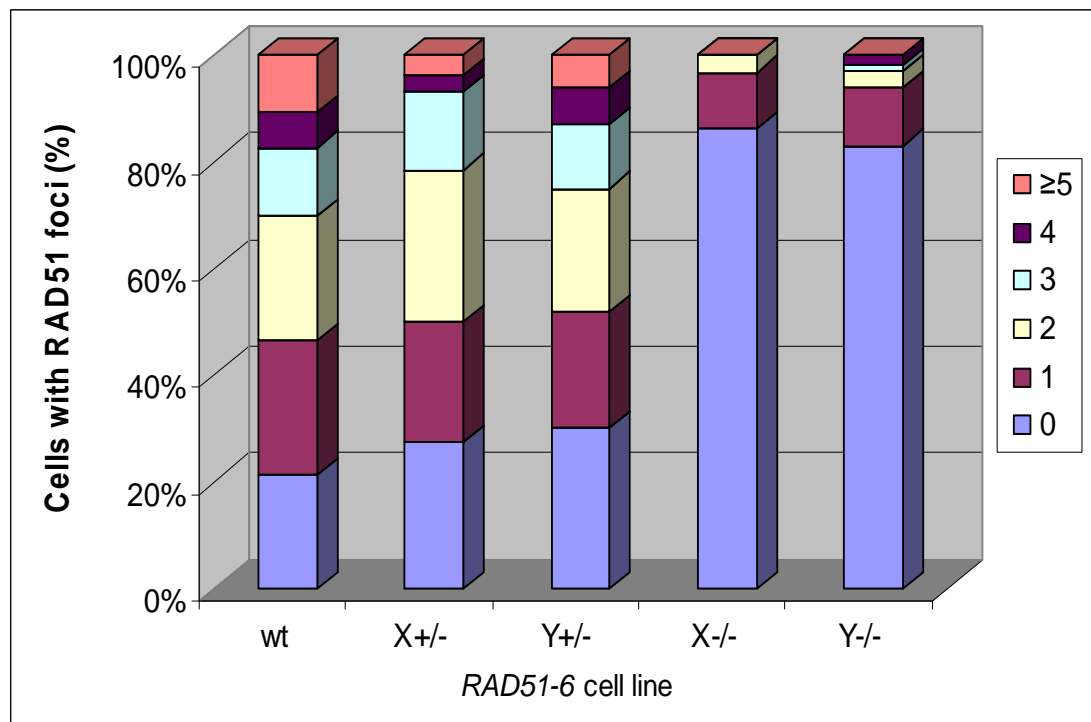
**Table 3-17: Percentage of *rad51-6*  $-/-$  mutant cells with a specific number of RAD51 foci.**

The percentage of cells with a specific number of the sub-nuclear RAD51 foci formed (0, 1, 2, 3, 4,  $\geq 5$ ) when the cell lines were treated with increasing concentrations of phleomycin,  $0.25 \mu\text{g.ml}^{-1}$  and  $1.0 \mu\text{g.ml}^{-1}$  is shown. The total indicates the number of cells counted. Wildtype strain 3174.2 (wt), two independent *RAD51-6* heterozygous mutants (X $+/$ -, Y $+/$ -), and two independent *rad51-6* homozygous mutants (X $-/-$ , Y $-/-$ ) were analysed.

	Concentration of phleomycin	Number of RAD51 foci						Cells Counted
		0	1	2	3	4	$\geq 5$	
wt	$0.25 \mu\text{g.ml}^{-1}$	32.0	43.0	14.0	9.0	1.0	5.0	100
	$1.0 \mu\text{g.ml}^{-1}$	21.3	25.3	23.1	12.9	6.7	10.7	225
X $+/$ -	$1.0 \mu\text{g.ml}^{-1}$	27.5	22.5	28.4	14.7	2.9	3.9	102
Y $+/$ -	$1.0 \mu\text{g.ml}^{-1}$	30.0	22.0	23.0	12.0	7.0	6.0	100
X $-/-$	$0.25 \mu\text{g.ml}^{-1}$	95.2	2.7	1.1	1.1	0.0	0.0	100
	$1.0 \mu\text{g.ml}^{-1}$	86.4	10.3	3.3	0.0	0.0	0.0	100
Y $-/-$	$0.25 \mu\text{g.ml}^{-1}$	96.6	1.7	0.9	0.9	0.0	0.0	100
	$1.0 \mu\text{g.ml}^{-1}$	83.0	11.0	3.0	1.0	2.0	0.0	100



In the *rad51-6*  $-/-$  mutants, 86-95 % of the cells did not display visible sub-nuclear RAD51 foci when treated with phleomycin either at a concentration of  $0.25 \mu\text{g.ml}^{-1}$  and  $1.0 \mu\text{g.ml}^{-1}$  (see Figure 3-36), compared with 21-32 % of wildtype cells. This result shows that cells lacking RAD51-6 are severely impaired in re-localisation of RAD51 in the nucleus, or in the maintenance of RAD51 foci once formed. It is, indeed, possible that the very small number of RAD51 foci scored are artefacts of punctate staining, which is seen in the absence of DNA damage (see Figure 3-35, and Section 3.4.4, Figure 3-15). This supports the theory that RAD51 re-localisation in response to DNA damage is aided by multiple factors, including RAD51-6, the other RAD51 paralogues and BRCA2 (Hartley, 2008; Proudfoot & McCulloch, 2005a). It also appears that the *rad51-6*  $-/-$  mutants are more severely impaired in this process than the *rad51-4*  $-/-$  mutants; this is considered further in Chapter 4, (Section 4.5).

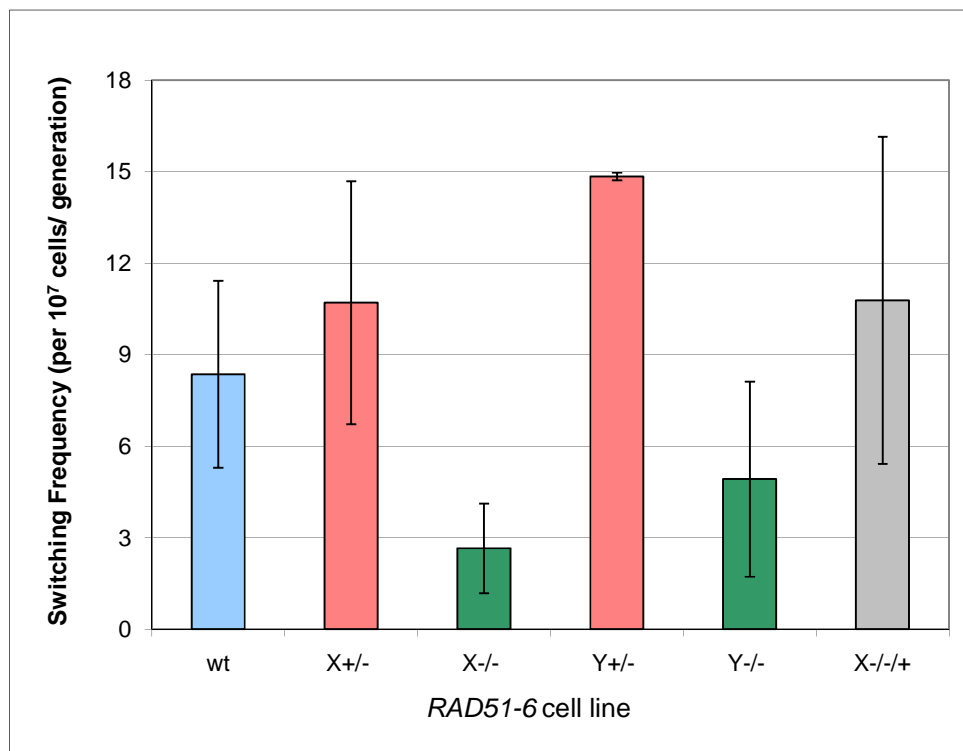


**Figure 3-36: Percentage of *rad51-6*  $-/-$  mutant cells with a specific number of RAD51 foci.**

The percentage of cells with a specific number of the sub-nuclear RAD51 foci formed (0, 1, 2, 3, 4,  $\geq 5$ ) when the cell lines were treated with  $1.0 \mu\text{g.ml}^{-1}$  of phleomycin for 18 hours. Wildtype strain 3174.2 (wt), two independent *RAD51-6* heterozygous mutants (X+/-, Y+/-), and two independent *rad51-6* homozygous mutants (X-/-, Y-/-) are shown.

### 3.7.5 Analysis of VSG switching in *RAD51-6* mutant cell lines

To examine the role of *RAD51-6* in VSG switching, the *rad51-6*  $-/-$  mutants were compared with wildtype, *RAD51-6*  $+/-$  and the *RAD51-6*  $-/-/+$  mutant cell lines in the VSG switching assay whose details are discussed in Section 3.4.5 and in Materials and Methods, Section 2.8.4. The results of this are shown in Figure 3-37 and Figure 3-38.



**Figure 3-37: The mean of the VSG switching frequencies in *RAD51-6* mutant cell lines.**

The graph shows the mean of the VSG switching frequencies of *RAD51-6* mutant cell lines. Wildtype strain 3174.2 (wt), two independent *RAD51-6* heterozygous (X $+/-$ , Y $+/-$ ), two independent *rad51-6* homozygous mutants (X $-/-$ , Y $-/-$ ) and one *RAD51-6* re-expressor (X $-/-/+$ ) are shown. The graph shows the means of triplicate data, with the 95 % confidence intervals indicated by the error bars.

Analysis of the VSG switching frequencies suggests that *rad51-6*  $-/-$  mutants have an impaired ability to undergo VSG switching as the VSG switching frequencies show a trend in which the *rad51-6*  $-/-$  switch at a lower frequency compared with the *RAD51-6*  $+/-$  cells from which they are derived, and compared with wildtype and re-expressor cells, see Figure 3-37. Unpaired two-tailed Student's *t*-tests were carried out on the data, and are shown in Table 3-18 and in Appendix 3, Table 14. In the independent clone X, the *RAD51-6* X $+/-$  had a VSG switching frequency of 10.7 ( $\pm$  3.9) switched variants per 10<sup>7</sup> cells per generation, which is significantly different from the related *rad51-6* X $-/-$ , which had a VSG switching frequency of 2.7 ( $\pm$  1.5) events per 10<sup>7</sup> cells per generation. The same was true for the other independent clone Y: *RAD51-6* Y $+/-$  had a VSG switching frequency

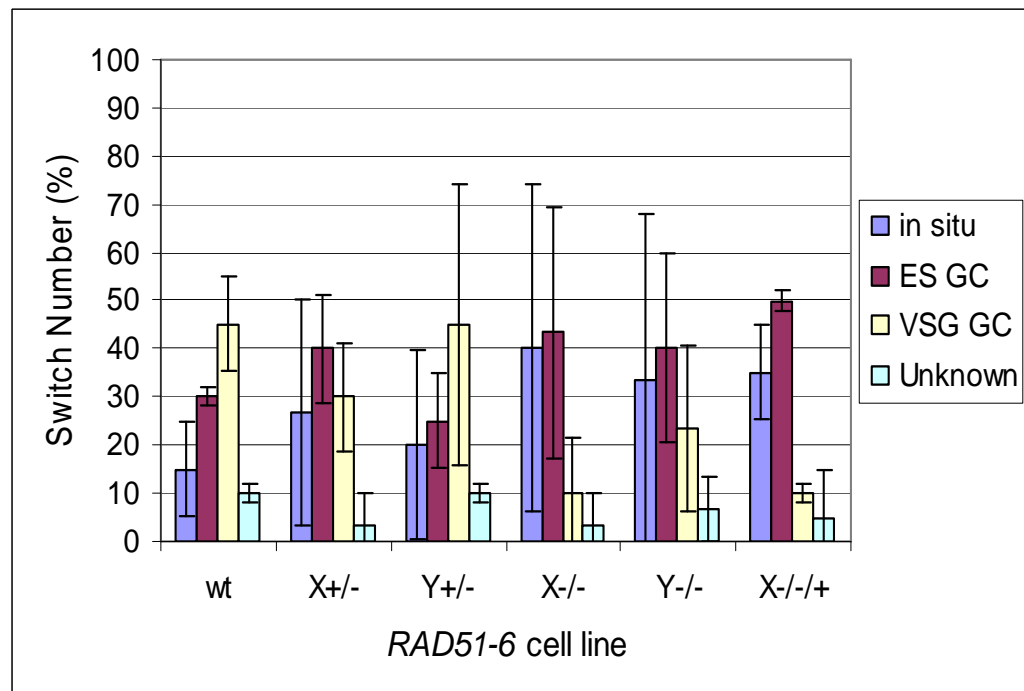
of 14.8 (+/- 0.1) switched variants per  $10^7$  cells per generation, which is significantly different from *rad51-6* Y<sup>-/-</sup>, which had a VSG switching frequency of 4.9 (+/- 3.2) switched variants per  $10^7$  cells per generation. Table 3-18 shows that there was also a significant difference between the switch frequencies of the wildtype compared with the *rad51-6* X<sup>-/-</sup> mutant and not the *rad51-6* Y<sup>-/-</sup> mutant. Similar to the results for the *rad51-4* <sup>-/-</sup> mutants, the large error bars (Figure 3-37) and lack of significant difference between the wildtype and both independent homologous mutants (Table 3-18) may suggest that RAD51-6 is not involved in VSG switching.

**Table 3-18: Statistical analysis of the VSG switching frequency of RAD51-6 mutant cell lines.**

Unpaired, two-tailed Student's *t*-tests were carried out to compare the mean of the triplicate VSG switching frequencies for the cell lines against the wildtype. A significant difference is shown by a *p*-value of  $p \leq 0.05$ . Not significantly different results are shown as "-". [NA is not applicable.] Wildtype strain 3174.2 (wt), two independent *RAD51-6* heterozygous (X<sup>+/-</sup>, Y<sup>+/-</sup>), two independent *rad51-6* homozygous mutants (X<sup>-/-</sup>, Y<sup>-/-</sup>) and one *RAD51-6* re-expressor (X<sup>-/-</sup>+) are shown.

	wt	X <sup>+/-</sup>	Y <sup>+/-</sup>	X <sup>-/-</sup>	Y <sup>-/-</sup>	X <sup>-/-</sup> +
wt	NA	-	-	0.0328	-	-

To examine this further, a selection of recovered switched clones from each cell line were analysed to differentiate the switching mechanisms that occurred, as described in Section 3.4.5 and Materials and Methods, Section 2.8.4. This analysis uses PCR amplification and drug sensitivity or resistance to predict the mechanism undertaken by the cells to switch their VSG coat. Given the involvement of RAD51-6 in recombination and repair, it might be predicted that the *rad51-6* <sup>-/-</sup> cells would not undergo gene conversion events to switch their VSG, since this mechanism relies on homologous recombination. Figure 3-36 show the switching mechanisms in the *rad51-6* <sup>-/-</sup> mutant cell lines, which demonstrates that these data did not fit this prediction, and expression site and VSG gene conversion reactions could be detected readily in both the *rad51-6* <sup>-/-</sup> mutants. It is also not clear whether or not *in situ* switching becomes more predominant, given the considerable variation of the data.



**Figure 3-38: Analysis of the switching mechanisms of the *RAD51-6* mutant cell lines.**

The graph shows the switching mechanism of *RAD51-6* mutant cell lines as a percentage of the total analysed. Wildtype strain 3174.2 (wt), two independent *RAD51-6* heterozygous mutants (X+/-, Y+/-), two independent *rad51-6* homozygous mutants (X-/-, Y-/-) and the *RAD51-6* re-expressor (X-/-/+) are shown. The assay allowed differentiation of *in situ* transcriptional switching (*in situ*), expression site gene conversion (ES GC) and VSG gene conversion (VSG GC). Switch reactions that could not be assigned in these categories are indicated as “unknown”. The graph shows the means of triplicate data, with the 95 % confidence intervals indicated by the error bars.

### 3.8 Summary

This chapter describes the generation and confirmation of gene knockouts of the two of remaining unstudied RAD51 paralogues, RAD51-4 and RAD51-6, in *T. brucei* blood stream form strain 3174.2. The *RAD51-4* and *RAD51-6* genes were disrupted by integration of antibiotic resistant cassettes in both alleles of each gene. These knockouts were generated using a reverse genetics approach, where the gene sequencing and phylogeny predicted the function of the *T. brucei* RAD51 paralogues having a role in one or more HR pathways. The phenotypic analyses described here confirmed this prediction and showed that RAD51-4 and RAD51-6 act in the HR pathway, with the two *rad51-4* -/- and *rad51-6* -/- mutants having a decrease in transformation efficiency. RAD51-4 and RAD51-6 may also affect antigenic variation, since the mean VSG switching frequencies were decreased in the homozygous compared to the heterozygous mutant cell lines. The *rad51-4* -/- and *rad51-6* -/- mutant cell lines also had a decreased growth rate and increased sensitivity to the DNA damaging agents, phleomycin and MMS. This analysis confirmed the prediction that RAD51-4 and RAD51-6 have roles in the RAD51-dependent pathway as the *rad51-4* -/- and *rad51-6* -/- mutants had a decrease in the amount of DNA-damage induced RAD51 foci formation. This result also confirms that the two paralogues

interact with RAD51 either directly or indirectly. As the *rad51-4*  $-/-$  and *rad51-6*  $-/-$  mutant cell lines are viable, this supports the belief that there are other RAD51-independent pathways involved in the double strand break repair of *T. brucei* DNA. The lack of functional repair pathways involving the RAD51 paralogues is therefore not completely detrimental to the cell.

It is interesting that the phenotype of the *rad51-4*  $-/-$  mutant cell lines appear to be less severely compromised compared to the *rad51-6*  $-/-$  mutants. This was especially true for the phleomycin sensitivities; where the  $IC_{50}$  value of the *rad51-4*  $-/-$  mutants is approximately double that of the *rad51-6*  $-/-$  mutants. This can also be seen when analysing the phleomycin-induced RAD51 foci formation percentages. Although the other phenotypic analyses are not truly comparable, the *rad51-4*  $-/-$  mutant cell lines are apparently affected to a lesser extent in comparison to the *rad51-6*  $-/-$  mutant cell lines when growth rate, HR rates and VSG switching frequencies are measured.

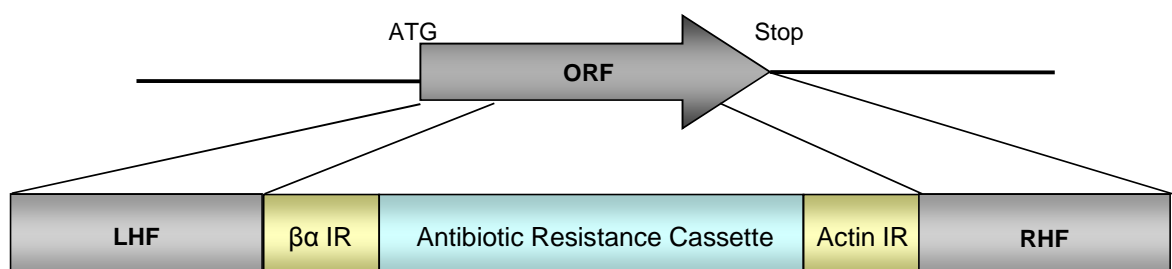
The impairment of growth rate, homologous recombination, DNA-damage induced repair and VSG switching of the *rad51-4*  $-/-$  and *rad51-6*  $-/-$  mutants are of a similar level compared with the other two RAD51 paralogue mutants, *rad51-3*  $-/-$  and *rad51-5*  $-/-$ , described previously by Proudfoot & McCulloch, 2005b. Chapter 4 of this thesis describes a more in depth analysis of the comparisons and contrasts all four *rad51* paralogue mutants alongside the wildtype and *rad51*  $-/-$  cell lines.

**Chapter 4. Comparing the impact  
of gene knockout of *RAD51* and  
the *RAD51* paralogues in *T.*  
*brucei***

## 4.1 Introduction

The aim of this chapter was to compare the severity of the different phenotypes of *RAD51* mutant and each of the four *RAD51* paralogue mutant cell lines. The rationale behind this was to attempt to determine if each factor contributes equally to DNA repair and recombination functions, and thereby ask if they act in a single or multiple pathways. To do this, all experiments were carried out, with the exception of the VSG switching assays, at the same time using the same equipment and methods. Therefore, this is a comprehensive re-analysis of the phenotypic data described in Chapter 3 and in Proudfoot & McCulloch (2005a), and allows direct comparison without the potential for experimental variation, due to unforeseen technical differences between experiments performed at different times and by different researchers.

The generation and confirmation of *rad51-3* *-/-* and *rad51-5* *-/-* cell lines is described in Proudfoot & McCulloch (2005a), and for *rad51-4* *-/-* and *rad51-6* *-/-* cell lines in the previous Chapter 3 of this thesis. The *rad51* *-/-* mutant was generated by Rebecca Barnes (unpublished). This was based on the constructs generated for the *RAD51* mutation by McCulloch & Barry (1999), but the phleomycin resistance construct was altered to provide blasticidin resistance. All the homozygous mutants were generated in an equivalent manner, with the disruption of the ORF by constructs that target the gene using left and right hand flanks of the start and end of the ORF respectively (Figure 4-1).



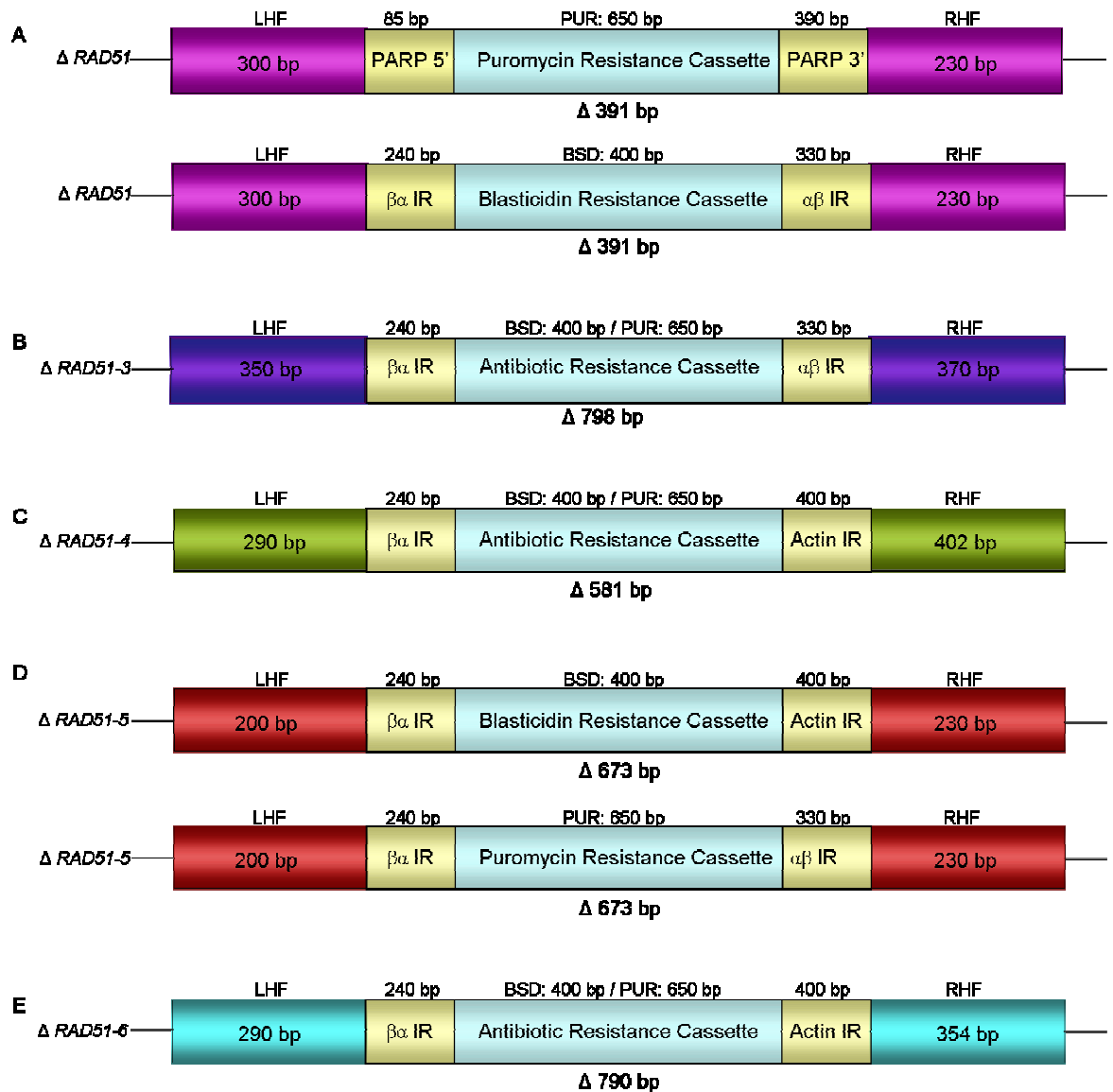
**Figure 4-1: Diagram showing a generic construct for the generation of *RAD51* and the *RAD51* paralogue mutants.**

The main ORF of *RAD51* and each of the *RAD51* paralogues were disrupted with an antibiotic resistance cassette, encoding resistance to blasticidin and puromycin. This was carried out through integration of the cassette by homologous recombination of targeting regions, the left hand flank (LHF) and right hand flank (RHF) of the gene.

In all cases the two alleles were disrupted using the blasticidin and puromycin antibiotic resistance cassettes. Use of these resistance cassettes, allowed analysis of phleomycin sensitivities of the cell lines. These constructs are shown in Figure 4-2 indicating the amount of ORF sequence deleted as a result of integration of the resistance cassette. Also shown are the intergenic regions flanking the resistance cassettes. For the *rad51* *-/-* mutant

generation, the  $\Delta RAD51::PUR$  construct had processing flanks derived from procyclin (McCulloch & Barry, 1999; Figure 4-2 A ). The  $\Delta RAD51-3::BSD$ ,  $\Delta RAD51-3::PUR$  and  $\Delta RAD51-5::PUR$  constructs differed in that both intergenic sequence (IR) flanking the antibiotic resistance ORFs were derived from tubulin,  $\beta\alpha$  and  $\alpha\beta$  IR (Proudfoot & McCulloch, 2005a; Figure 4-2 B, D). All the other constructs,  $\Delta RAD51-4::BSD/PUR$ ,  $\Delta RAD51-5::BSD$  and  $\Delta RAD51-6::BSD/PUR$  had one intergenic region derived from the tubulin and the other from actin (Figure 4-2 C, D, E). All these homozygous mutants were generated in the same *T. brucei* strain 3174.2 in which the VSG switching frequencies and mechanisms were analysed (McCulloch *et al.*, 1997; McCulloch & Barry, 1999).





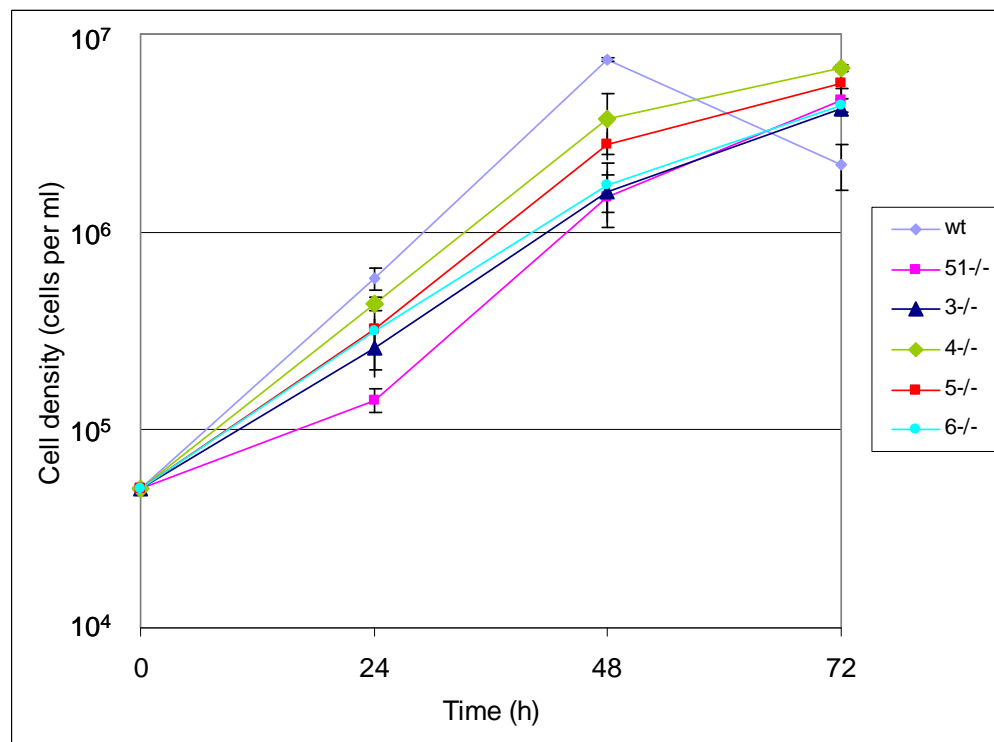
**Figure 4-2: Diagram showing the knockout constructs for the generation of *rad51* and the *RAD51* paralogue mutant cell lines.**

The main ORF of *RAD51* and each *RAD51* paralogue was disrupted with an antibiotic resistance cassette, encoding resistance to blasticidin (*BSD*) and puromycin (*PUR*). This was carried out through integration of the cassette by homologous recombination of targeting regions, the left hand flank (*LHF*) and right hand flank (*RHF*) of the gene, and the sequence sizes are shown in base pairs (bp). The diagram shows the constructs used: **Figure 4-2A:**  $\Delta RAD51::PUR$ ,  $\Delta RAD51::BSD$ , **Figure 4-2B:**  $\Delta RAD51-3::BSD/PUR$ , **Figure 4-2C:**  $\Delta RAD51-4::BSD/PUR$ , **Figure 4-2D:**  $\Delta RAD51-5::BSD$ ,  $\Delta RAD51-5::PUR$ , and **Figure 4-2E:**  $\Delta RAD51-6::BSD/PUR$ , each indicating the amount of the ORF deleted by integration of the construct,  $\Delta 391$  bp,  $\Delta 798$  bp,  $\Delta 581$  bp,  $\Delta 673$  bp, and  $\Delta 790$  bp respectively. The intergenic regions flanking the antibiotic resistance genes are indicated as follows:  $\beta\alpha$  and  $\alpha\beta$  tubulin ( $\beta\alpha$  IR and  $\alpha\beta$  IR), actin intergenic region (Actin IR) and procyclin intergenic regions (PARP 5' and PARP 3').

The *rad51*  $-/-$  mutant cells and the equivalent homozygous *RAD51* paralogue mutants, *rad51-3*  $-/-$ , *rad51-4*  $-/-$ , *rad51-5*  $-/-$  and *rad51-6*  $-/-$ , were all viable. Thus, phenotypic analysis of each cell line is possible, as described in this chapter. These results will be considered in comparison with single *RAD51* paralogue gene mutation analysis in other species, in particular vertebrate cells where the proteins appear to adopt overlapping, yet distinct functions (Liu *et al.*, 2004; Sigurdsson *et al.*, 2001; Takata *et al.*, 2000; Takata *et al.*, 2001; Tarsounas *et al.*, 2004b).

## 4.2 Analysis of the *in vitro* growth rates of *RAD51* and *RAD51* paralogue mutants

Previous work has shown that homozygous mutants of *rad51* and each of the *RAD51* paralogues had a reduced *in vitro* growth rate (McCulloch & Barry, 1999; Proudfoot & McCulloch, 2005a). In some cases, this was also seen *in vivo* (McCulloch & Barry, 1999). To compare the extent of this impairment, growth curves and population doubling time were calculated (this was performed as before; see Materials and Methods, Section 2.8). Figure 4-3 and Table 4-1 confirms that the *RAD51* and each *RAD51* paralogue homozygous mutant, (*rad51-3* *-/-*, *rad51-4* *-/-*, *rad51-5* *-/-* and *rad51-6* *-/-*) have a growth impairment compared with that of the wildtype. However, it is notable that, *rad51-4* *-/-* and *rad51-5* *-/-* mutant growth rates (green and red lines of Figure 4-3) appear to be faster than that of the other *rad51* *-/-*, *rad51-3* *-/-* and *rad51-6* *-/-* (see Table 4-1). The same trend was observed previously, in a more limited sample: Proudfoot & McCulloch, (2005a) observed that *rad51-5* *-/-* mutants appeared to grow quicker than the *rad51-3* *-/-* mutants. The implications of this are unknown because the basis for slow growth in the mutants is unclear, but it may suggest that these two *RAD51* paralogue proteins, *RAD51-4* and *RAD51-5*, may have less central functions in cell viability or cell cycle progression, perhaps due to having less crucial roles in DNA repair. It may be, for instance, that *RAD51-4* and *RAD51-5* are not as integral to DNA repair as *RAD51* and the other *RAD51* paralogues, or that they have a role or roles that can be performed by some other protein or protein complexes. It may be that their role is to aid or stabilise the other *RAD51* paralogues that are more centrally involved in repair, which has been suggested for some members of the mammalian *RAD51* family (Lio *et al.*, 2004).



**Figure 4-3: Analysis of the growth of *RAD51* and *RAD51* paralogue mutants *in vitro***

A semi logarithmic plot of the cell densities at fixed time points, 24 h, 48 h, and 72 h is shown. This plot shows the growth of *rad51* <sup>-/-</sup> (51<sup>-/-</sup>), *rad51-3* <sup>-/-</sup> (3<sup>-/-</sup>), *rad51-4* <sup>-/-</sup> (4<sup>-/-</sup>), *rad51-5* <sup>-/-</sup> (5<sup>-/-</sup>), and *rad51-6* <sup>-/-</sup> (6<sup>-/-</sup>) mutants compared with the wildtype strain 3174.2 (wt). The means of three independent experiments are shown, with the 95 % confidence intervals indicated by error bars.

**Table 4-1: *in vitro* population doubling times for the homozygous mutants of *RAD51* and the *RAD51* paralogues.**

The population doubling time of each homozygous mutant, *rad51* <sup>-/-</sup>, *rad51-3* <sup>-/-</sup>, *rad51-4* <sup>-/-</sup>, *rad51-5* <sup>-/-</sup>, and *rad51-6* <sup>-/-</sup>, are displayed in hours. These data are compared with the population doubling time for the wildtype strain 3174 (wt).

wt	<i>rad51</i> -/-	<i>rad51-3</i> -/-	<i>rad51-4</i> -/-	<i>rad51-5</i> -/-	<i>rad51-6</i> -/-
6.69 (+/- 0.019)	9.99 (+/- 1.0)	9.80 (+/- 0.89)	7.87 (+/- 0.73)	8.37 (+/- 0.31)	9.52 (+/- 0.79)

**Table 4-2: Statistical analysis of the *in vitro* population doubling times for the homozygous mutants of *RAD51* and the *RAD51* paralogues.**

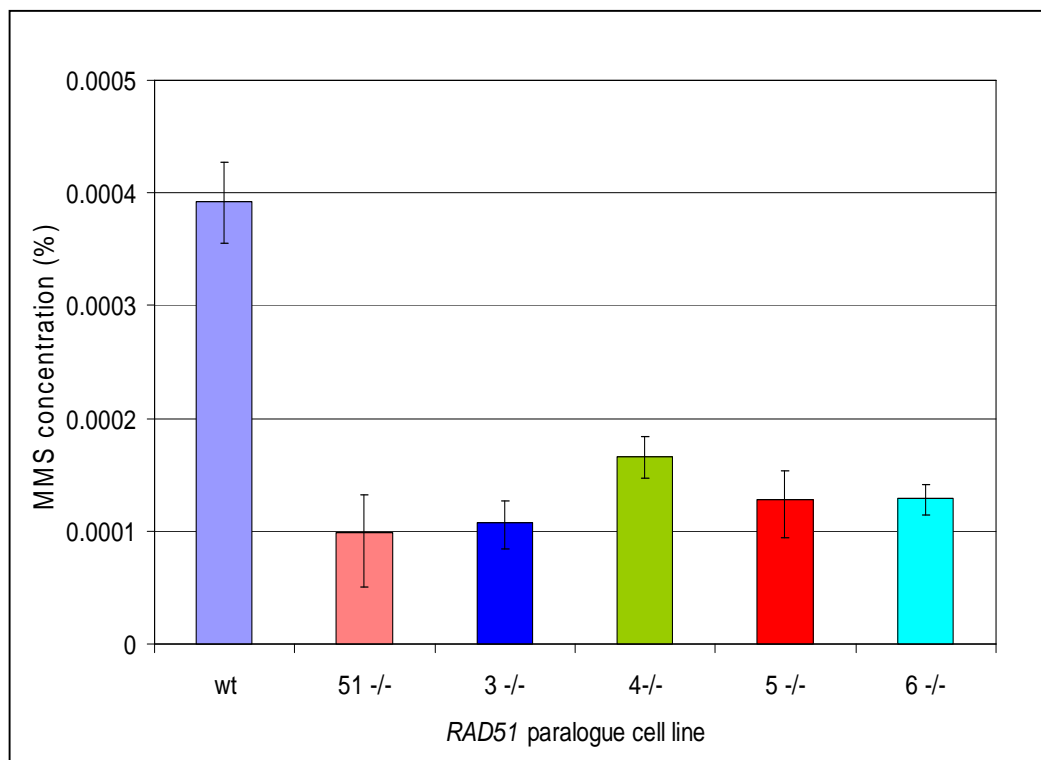
Unpaired, two-tailed Student's *t*-tests were carried out to compare the mean of the triplicate growth rates for the wildtype (wt) and *rad51* *-/-*, *rad51-3* *-/-*, *rad51-4* *-/-*, *rad51-5* *-/-* and *rad51-6* *-/-* mutants. A significant difference between the means is shown by a *p*-value of  $p \leq 0.05$ . Not significantly different results are shown as “-“. NA is not applicable.

	wt	<i>rad51</i> <i>-/-</i>	<i>rad51-3</i> <i>-/-</i>	<i>rad51-4</i> <i>-/-</i>	<i>rad51-5</i> <i>-/-</i>	<i>rad51-6</i> <i>-/-</i>
wt	NA	0.0030	0.0025	0.0338	0.0005	0.0022
<i>rad51</i> <i>-/-</i>	0.0030	NA	-	0.0290	0.0395	-
<i>rad51-3</i> <i>-/-</i>	0.0025	-	NA	0.0309	0.0419	-
<i>rad51-4</i> <i>-/-</i>	0.0338	0.0290	0.0309	NA	-	0.0400
<i>rad51-5</i> <i>-/-</i>	0.0005	0.0395	0.0419	-	NA	-
<i>rad51-6</i> <i>-/-</i>	0.0022	-	-	0.0400	-	NA

It has been observed in chicken DT40 cells, that the growth rates of the Rad51 paralogue mutants, *rad51b*, *rad51c*, *rad51d*, *xrcc2* and *xrcc3*, are significantly lower when compared with the wildtype cells (Takata *et al.*, 2000; Takata *et al.*, 2001). The authors noted that the cell cycle profiles were the same for the wildtype and *rad51* paralogue mutants, and suggested that the lower growth rates were due to the high percentage (~20-30 %) of cell death in the *rad51* paralogue mutant cell lines (Takata *et al.*, 2000; Takata *et al.*, 2001). This has also been suggested to be true for *rad51-3* *-/-* and *rad51-5* *-/-* mutants in *T. brucei* (Proudfoot & McCulloch, 2005a), but the cell cycle profiles of *rad51-4* *-/-* and *rad51-6* *-/-* mutant cell lines were not examined. It is also noteworthy that the *rad51* paralogue mutants in the DT40 cells display chromosomal aberrations, as do the *T. brucei brca2* *-/-* cells, and this may be the underlying cause of cell death in the population and reduced growth rates (Hartley & McCulloch, 2008; Liu *et al.*, 1998; Takata *et al.*, 2001). However, the reduced growth rate in *T. brucei brca2* *-/-* is partly due to defects in the nuclear DNA replication or cell division (Hartley, PhD thesis), although this has not been observed in *rad51* *-/-* nor in the *RAD51* paralogue homozygous mutants. Nevertheless, further reasons why the *rad51* paralogue mutants in all species studied have reduced growth rates could yet be uncovered. In addition, determining whether *T. brucei rad51* or *RAD51* paralogue homozygous mutants display chromosomal aberrations has not been tested.

### 4.3 Analysis of the DNA-damage sensitivity of *RAD51* and *RAD51* paralogue mutants

To analyse whether mutations in *RAD51* and the *RAD51* paralogues lead to the same level of DNA-damage sensitivity,  $IC_{50}$  values for two DNA-damaging agents, methyl methanesulfonate (MMS) and phleomycin, were calculated. As previously described in Chapter 3 (Section 3.4.2.2), each homozygous mutant cell line was treated with serial dilutions of the DNA-damaging agents and the resulting proliferation was measured by resazurin, which was used as an indicator of growth. Both DNA-damaging agents have been observed to directly or indirectly induce double strand breaks (Claussen & Long, 1999; Giloni *et al.*, 1981; Lundin *et al.*, 2005; Wyatt & Pittman, 2006). Figures 4-4 and 4-5 confirms the previously described data that the *rad51*  $-/-$  mutant and each *RAD51* paralogue homozygous mutant are more sensitive than wildtype cells to both DNA-damaging agents. The significantly different results are indicated in Table 4-3 and Table 4-4. The mutants are 2.0 - 3.5 fold more sensitive to MMS and 2.5 – 5.0 fold more sensitive to phleomycin compared with treated wildtype cells. However, the *rad51-4*  $-/-$  mutant appears to be more resistant to phleomycin treatment than the *rad51*  $-/-$  mutant and the other *RAD51* paralogue homozygous mutants. As for the growth rate difference, the reason for this greater resistance is unknown but appears to be reflected by analysing the formation of *RAD51* foci in response to DNA-damage by phleomycin (see Section 4.5 below). Arguably, these differences, in growth rates and DNA-damage sensitivities, are the result of distinct functions for *RAD51-4* in the DNA repair pathways used by *T. brucei*. However, it is not clear why the *rad51-4*  $-/-$  cells are not also more resistant to MMS, which also causes DNA breaks. Furthermore, *rad51-5*  $-/-$  cells appeared not to have as severe growth defects as *rad51*  $-/-$ , *rad51-3*  $-/-$ , and *rad51-6*  $-/-$  mutants, but displays essentially equivalent levels of DNA damage sensitivity.



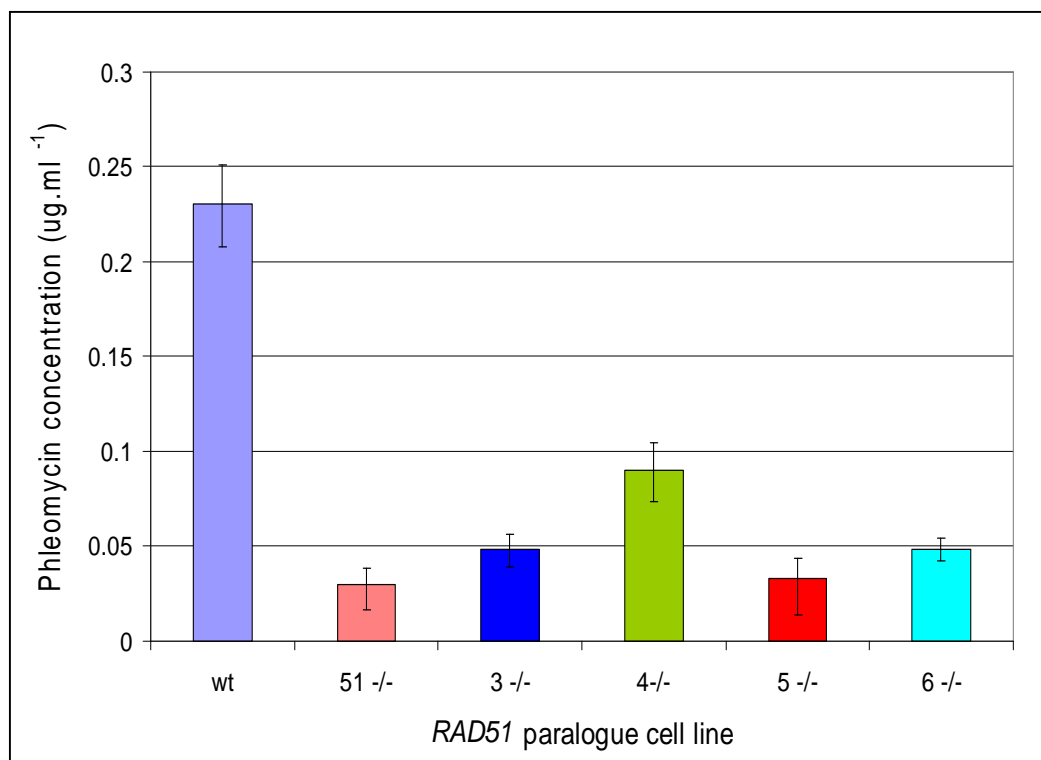
**Figure 4-4: IC<sub>50</sub> values of MMS for the homozygous mutants of *RAD51* and the *RAD51* paralogues.**

Each cell line was treated with serial dilutions of MMS and the resulting proliferation was measured by resazurin, which was used as an indicator of growth. The IC<sub>50</sub> values are shown for the following cell lines: *rad51* -/- (51-/-), *rad51-3* -/- (3-/-), *rad51-4* -/- (4-/-), *rad51-5* -/- (5-/-), and *rad51-6* -/- (6-/-), and compared with wildtype 3174 strain (wt). The values are the mean of three independent experiments, and 95 % confidence intervals are indicated by the error bars.

**Table 4-3: Statistical analysis of the IC<sub>50</sub> values of MMS for the homozygous mutants of *RAD51* and the *RAD51* paralogues.**

Unpaired, two-tailed Student's *t*-tests were carried out to compare the IC<sub>50</sub> values of MMS for the wildtype (wt) and *rad51* -/-, *rad51-3* -/-, *rad51-4* -/-, *rad51-5* -/- and *rad51-6* -/- mutant cell lines. A significant difference between the means is shown by a *p*-value of  $p \leq 0.05$ . Not significantly different results are shown as “-“. NA is not applicable.

	wt	<i>rad51</i> -/-	<i>rad51-3</i> -/-	<i>rad51-4</i> -/-	<i>rad51-5</i> -/-	<i>rad51-6</i> -/-
wt	NA	0.0092	0.0098	0.0195	0.0130	0.0127
<i>rad51</i> -/-	0.0092	NA	-	0.0364	-	-
<i>rad51-3</i> -/-	0.0098	-	NA	0.0453	-	-
<i>rad51-4</i> -/-	0.0195	0.0364	0.0453	NA	-	-
<i>rad51-5</i> -/-	0.0112	-	-	-	NA	-
<i>rad51-6</i> -/-	0.0127	-	-	-	-	NA



**Figure 4-5: IC<sub>50</sub> values of phleomycin for the homozygous mutants of *RAD51* and the *RAD51* paralogues.**

Each cell line was treated with serial dilutions of phleomycin and the resulting proliferation was measured by resazurin, which was used as an indicator of growth. The IC<sub>50</sub> values are shown for the following cell lines: *rad51* -/- (51-/-), *rad51-3* -/- (3-/-), *rad51-4* -/- (4-/-), *rad51-5* -/- (5-/-), and *rad51-6* -/- (6-/-), and compared with wildtype 3174 strain (wt). The values are the mean of three independent experiments, and 95 % confidence intervals are indicated by the error bars.

**Table 4-4: Statistical analysis of the IC<sub>50</sub> values of phleomycin for the homozygous mutants of *RAD51* and the *RAD51* paralogues.**

Unpaired, two-tailed Student's *t*-tests were carried out to compare the IC<sub>50</sub> values of phleomycin for the wildtype (wt) and *rad51* -/-, *rad51-3* -/-, *rad51-4* -/-, *rad51-5* -/- and *rad51-6* -/- mutant cell lines. A significant difference between the means is shown by a *p*-value of  $p \leq 0.05$ . Not significantly different results are shown as "-". NA is not applicable.

	wt	<i>rad51</i> -/-	<i>rad51-3</i> -/-	<i>rad51-4</i> -/-	<i>rad51-5</i> -/-	<i>rad51-6</i> -/-
wt	NA	0.0173	0.0224	0.0372	0.0181	0.0211
<i>rad51</i> -/-	0.0173	NA	-	0.0140	-	0.0459
<i>rad51-3</i> -/-	0.0224	-	NA	0.0495	-	-
<i>rad51-4</i> -/-	0.0372	0.0140	0.0495	NA	0.0174	0.0277
<i>rad51-5</i> -/-	0.0181	-	-	0.0174	NA	-
<i>rad51-6</i> -/-	0.0211	0.0459	-	0.0277	0.0438	NA

The mammalian V79 and CHO Chinese hamster cell lines *irs1* and *irs1SF* (defective in the mammalian Rad51 paralogues, XRCC2 and XRCC3, respectively) were first described to be moderately sensitive to x-rays and gamma radiation, and extremely sensitive to DNA cross-linking agents (Liu et al., 2007; Tebbs et al., 1995). Later, it was shown that all the Rad51 paralogue mutants in chicken DT40 cell lines were similarly moderately sensitive to gamma radiation, and extremely sensitive to DNA cross-linking agents, as assayed by clonal survival (Takata et al., 2000; Takata et al., 2001). In *S. cerevisiae*, *rad55* and *rad57* null mutants were more sensitive to x-rays at 23 °C compared with wildtype at the same temperature (Hays et al., 1995). Complementation studies using these mutant cell lines have shown that overexpression of other DNA repair proteins such as Rad51 and Rad52, overcome these phenotypes (Hays et al., 1995; Johnson & Symington, 1995). Both Rad52 and Rad55/57 are required to promote Rad51 filament assembly in irradiated mitotic cells, whereas either Rad52 or Rad55/57 alone are insufficient (Gasior et al., 2001). Also addition of corresponding human cDNA to the chicken *rad51* paralogue mutants restores some if not all of the DNA-damage resistance (Takata et al., 2001). These findings suggest that the Rad51 paralogues support and facilitate the function of Rad51.

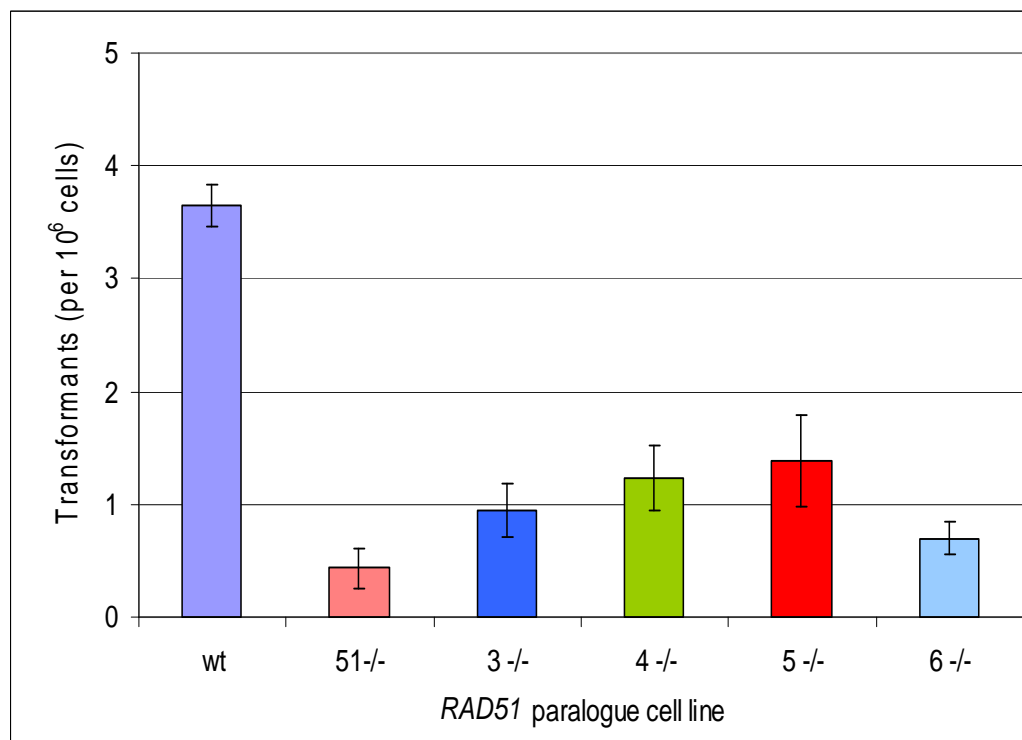
There is diversity among the different DNA-damaging agents used in the experiments described above, and there is a precedent for the eukaryotic Rad51 paralogue mutants to respond differently to various DNA-damaging agents (Game & Mortimer, 1974; Jones et al., 1987; Jones et al., 1988). The *T. brucei* RAD51 paralogues have also been shown to act in response to DNA-damaging agents, MMS and phleomycin, which most likely yield a similar profile of DNA breaks. The effect of other types of DNA-damaging agents on the *RAD51* paralogue mutants has not been assessed. For instance, there is no information on how the *T. brucei* *RAD51* paralogue homozygous mutants would react to DNA-crosslinking agents such as cisplatin or to replication stalling compounds such as hydroxyurea. Analysis of DNA-damaging agents with differing modes of action may reveal more information on the roles of the RAD51 paralogues.



## 4.4 Analysis of the recombination efficiency of *RAD51* and the *RAD51* paralogue mutants

To analyse the ability of the different mutants to undergo homologous recombination (HR), quantification of the cells' abilities to integrate a transformed DNA construct that is targeted to the  $\beta\alpha$  tubulin array was performed (see Chapter 6 for a more detailed discussion). Briefly, to assay for transformation efficiency, a linearised construct containing the drug resistant gene for phleomycin resistance, flanked by  $\beta\alpha$  and  $\alpha\beta$  intergenic regions, was transformed into the wildtype and each of homozygous mutant cell lines (see Materials and Methods, Section 2.8.3). If HR can occur, the  $\beta\alpha$  and  $\alpha\beta$  intergenic flanks allow the resistance cassette to integrate into the  $\beta\alpha$  tubulin array, and following expression of the cassette the cells become resistant to phleomycin. This is potentially a highly variable assay, and can depend on many factors, including quantity and purity of the linearised DNA, density and growth conditions of the *T. brucei* cells, efficiency of electroporation, and concentration of phleomycin used for selection. For this reason, it was important to analyse the cells lines at the same time with the same reagents, in order to ascertain if differences or similarities in recombination efficiency are observed.

Figure 4-6 confirms all previous data suggesting that the *rad51* and each of the *RAD51* paralogue homozygous mutants have a decrease in transformation efficiency compared to wildtype cells. It is notable, however, that the rate of transformation in this data set is approximately 4 fold higher for all the cell lines, than observed previously in Chapter 3 (Section 3.4.3 and Section 3.7.3). This illustrates the importance of comparing results within a single experiment. Nevertheless, the *rad51* *-/-* mutant appears to have a greater decrease in its ability to undergo HR than any of the *RAD51* paralogue homozygous mutants, and may reflect the central importance of RAD51 directly in the pathway, compared with the putative accessory functions of the RAD51 paralogues. Of the four *RAD51* paralogue mutants, the *rad51-6* *-/-* mutant appears to have the greatest impairment, which is nearly comparable with that of the *rad51* *-/-* cells and the transformation efficiencies of the two cell lines, *rad51* *-/-* and *rad51-6* *-/-*, are not significantly different from one another (see Table 4-5). This result may imply that RAD51-6 has a more fundamental role in this form of homologous recombination compared with the other RAD51 paralogues, though further analysis would be needed to support this suggestion. It should be highlighted that mutation of either *RAD51* or any of the *RAD51* paralogues, did not result in total obliteration of HR, suggesting that other pathways can act to recombine DNA, assuming that all these proteins contribute to RAD51-directed recombination.



**Figure 4-6: Transformation assay for *RAD51* and the *RAD51* paralogue mutants**

To assay for recombination rate, a transformation efficiency assay was used. Each cell line is indicated as follows: Wildtype strain 3174.2 (wt), *rad51*  $-/-$  (51-/-), *rad51-3*  $-/-$  (3-/-), *rad51-4*  $-/-$  (4-/-), *rad51-5*  $-/-$  (5-/-), and *rad51-6*  $-/-$  (6-/-). The values plotted are mean numbers of phleomycin resistant transformants per 10<sup>6</sup> cells transformed. These experiments were done in triplicate with the same linearised DNA, and error bars of the 95 % confidence intervals are indicated.

**Table 4-5: Statistical analysis of the transformation efficiency for the homozygous mutants of *RAD51* and the *RAD51* paralagues.**

Unpaired, two-tailed Student's *t*-tests were carried out to compare the transformation efficiency for the wildtype (wt) and *rad51*  $-/-$ , *rad51-3*  $-/-$ , *rad51-4*  $-/-$ , *rad51-5*  $-/-$  and *rad51-6*  $-/-$  mutant cell lines. A significant difference between the means is shown by a *p*-value of  $p \leq 0.05$ . Not significantly different results are shown as “-“. NA is not applicable.

	wt	<i>rad51</i> -/-	<i>rad51-3</i> -/-	<i>rad51-4</i> -/-	<i>rad51-5</i> -/-	<i>rad51-6</i> -/-
wt	NA	0.0000	0.0000	0.0000	0.0000	0.0000
<i>rad51</i> -/-	0.0000	NA	0.0073	0.0008	0.0015	-
<i>rad51-3</i> -/-	0.0000	0.0073	NA	-	-	-
<i>rad51-4</i> -/-	0.0000	0.0008	-	NA	-	0.0189
<i>rad51-5</i> -/-	0.0000	0.0015	-	-	NA	0.0260
<i>rad51-6</i> -/-	0.0000	-	-	0.0189	0.0260	NA

It has been shown that XRCC2, XRCC3 and Rad51C in mammalian cells have a role in HR (Johnson *et al.*, 1999; Lio *et al.*, 2004; Pierce *et al.*, 1999). These studies used assay systems (such as those discussed in Chapter 6) where HR is measured using markers located within or between chromosomes, as opposed to the indirect transformation method described here for *T. brucei*. Despite the different methods used, both the *T. brucei* and mammalian Rad51 paralogues have been shown to have a role in HR, with mutants of the Rad51 paralogues having a decrease in the rate of recombination. This was also the case with the *T. brucei brca2* *-/-* mutants (Hartley & McCulloch, 2008). These analyses confirm that the RAD51 paralogues have a role in HR along with RAD51 and BRCA2.

## 4.5 Analysis of the role of the RAD51 paralogues in RAD51 foci formation

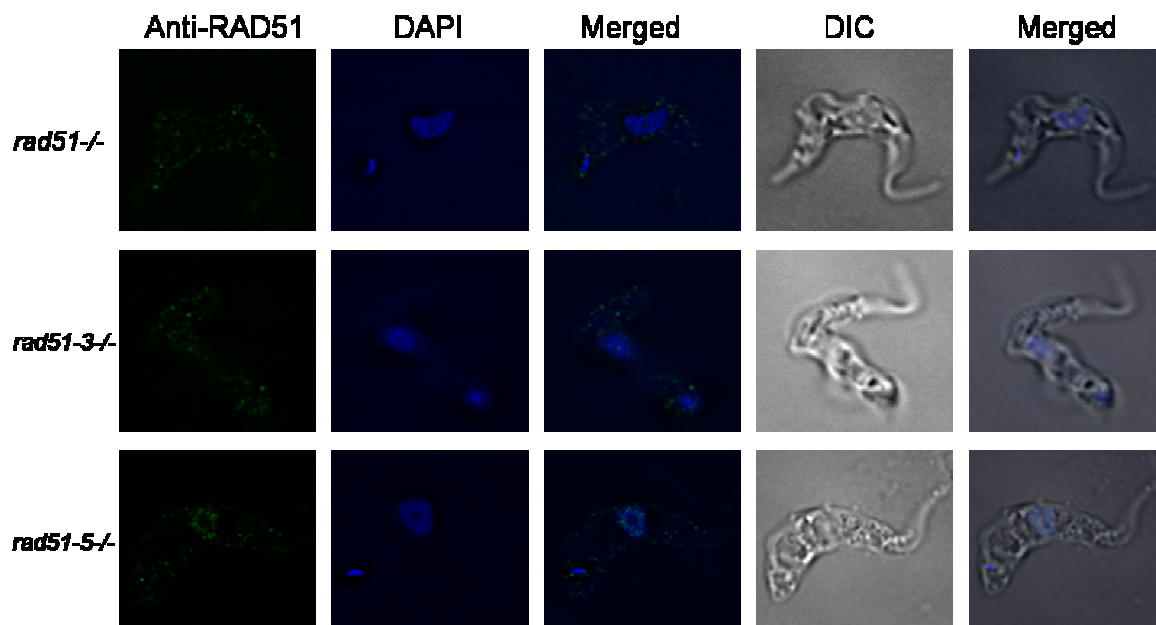
It has been shown that RAD51 forms discrete foci in sub-nuclear complexes in response to different forms of DNA-damage (Bishop, 1994; Haaf *et al.*, 1995; Tarsounas *et al.*, 2004a). It has similarly been shown that following phleomycin damage of *T. brucei* cells, RAD51 forms distinct foci (Glover *et al.*, 2008; Hartley & McCulloch, 2008; Proudfoot & McCulloch, 2005a). RAD51 foci are most probably formed at the sites of DNA damage and may represent regions of repair of DNA lesions (Raderschall *et al.*, 1999; Tarsounas *et al.*, 2004a; Tashiro *et al.*, 2000). As discussed in previous publications and Chapter 3 (Section 3.4.4), the number of RAD51 foci formed in the *T. brucei* damaged cells varies and can be counted using indirect immuno-fluorescence using anti-RAD51 anti-sera. However, although it has been shown that the *rad51-4* *-/-* and *rad51-6* *-/-* mutants have a reduced ability to form RAD51 foci in response to phleomycin (Chapter 3, Section 3.4.4 and 3.7.4), similar to that described previously for *rad51-3* *-/-*, *rad51-5* *-/-*, and *brca2* *-/-*, the method used here (Materials and Methods, Section 2.9). As a result, comparable quantification of the RAD51 foci in all the *RAD51* paralogue homozygous mutants was required, in order to determine if each factor contributes equally in this process.

To examine the RAD51 foci formation response to phleomycin in more detail,  $1 \times 10^6$  cells per ml of each *RAD51* paralogue homozygous mutant were grown in the presence of 0.0  $\mu\text{l}.\text{ml}^{-1}$ , 0.25  $\mu\text{l}.\text{ml}^{-1}$  and 1.0  $\mu\text{l}.\text{ml}^{-1}$  of phleomycin for 18 hours (see Materials & Methods, Section 2.9). Briefly, the treated cells were washed in PBS, and fixed using 1 % formaldehyde/PBS. After fixing, the cells were washed again with PBS and permeabilised with 1 % Triton X-100 and spread on to a glass slide. The cells were probed with rabbit anti-RAD51 anti-serum at a dilution of 1:500 and detected with goat anti rabbit-conjugated with SFX (fluorescein, succinimidyl ester). The addition of mounting solution containing

DAPI allowed detection of the nucleus and kinetoplast. Once the slide had dried completely, the number of RAD51 foci per cell was scored microscopically. For comparison, this was also performed for wildtype and non-treated cell lines.

In the cells not treated with phleomycin minimal RAD51 foci formation (1-3 %) were observed, as no DNA damage was induced (data not shown). Phleomycin-treated *rad51*  $-/-$  mutants also did not produce any RAD51 foci, because RAD51 is absent in these cells, see Figure 4-7. Table 4-6 and Figure 4-8 show the percentage of RAD51 paralogue mutant cells with a specific number of RAD51 foci observed when treated with phleomycin. The data for the *rad51-4*  $-/-$  and *rad51-6*  $-/-$  mutants are those described in Chapter 3 (Sections 3.4.4 and 3.7.4). In the treated *rad51-3*  $-/-$  and *rad51-5*  $-/-$  mutants, there were few observable RAD51 foci (Table 4-6 and Figure 4-8) and the percentage of cells that produced RAD51 foci is much lower when compared with the treated wildtype cells. The number of treated *rad51-3*  $-/-$  and *rad51-5*  $-/-$  cells with distinct RAD51 foci is lower in this analysis than in the study of Proudfoot & McCulloch (2005a). When treated with a phleomycin concentration of  $0.25 \mu\text{g.ml}^{-1}$ , 93 % of *rad51-3*  $-/-$  and 82 % of *rad51-5*  $-/-$  mutant cells had no observable RAD51 foci using the technique described above (Table 4-6), while 89 % and 83 % of cells treated with  $1.0 \mu\text{g.ml}^{-1}$  phleomycin had no foci respectively. At a near equivalent phleomycin concentration, Proudfoot & McCulloch (2005a) found a higher percentage of cells with RAD51 foci. 60-80 % and 52-64 % of *rad51-3*  $-/-$  mutant cells did not show RAD51 foci formation when treated with  $0.3 \mu\text{g.ml}^{-1}$  and  $1.5 \mu\text{g.ml}^{-1}$  phleomycin respectively. A similar result was seen when *rad51-5*  $-/-$  mutant cells were treated with  $0.5 \mu\text{g.ml}^{-1}$  and  $1.5 \mu\text{g.ml}^{-1}$  phleomycin, as 64-67 % and 60-64 % of cells respectively, did not have detectable RAD51 foci. It was apparent that the amount of overall anti-RAD51 staining was greater in the previous study than in this, perhaps due to greater amounts of non-specific antibody binding, and therefore one explanation is that a greater number of “false” foci were scored.

A striking observation from the data presented in Figure 4-8 is that the *rad51-4*  $-/-$  mutant appears less impaired in its ability to support relocalisation of RAD51 into foci than the other three RAD51 paralogue mutants. RAD51 foci formation in the treated *rad51-3*  $-/-$ , *rad51-5*  $-/-$ , and *rad51-6*  $-/-$  cell lines appear to be broadly similar, with very few distinct RAD51 foci observed. This appears to be consistent with the findings in Section 4.3 above, which shows that the *rad51-4*  $-/-$  mutants are less sensitive to phleomycin DNA damage than in the other mutants. This appears to suggest a less central role for RAD51-4 in this form of DNA repair.



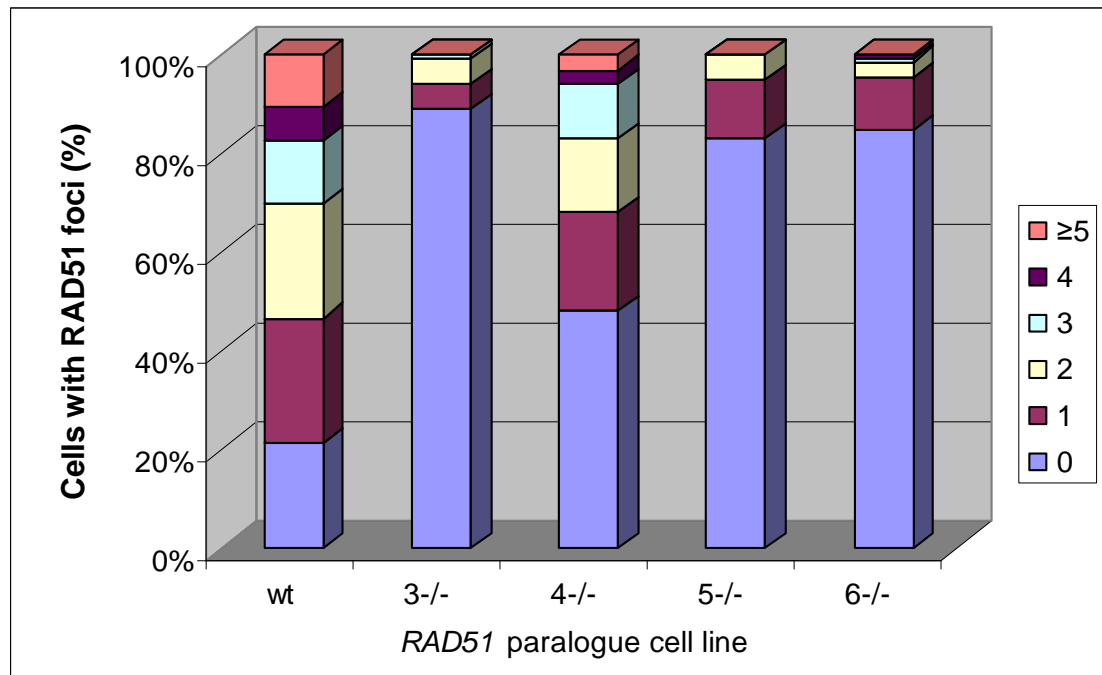
**Figure 4-7: Detection of the RAD51 foci by indirect immuno-fluorescence.**

The first row of pictures shows an example of a *rad51* <sup>-/-</sup> cell treated with 1.0 µg.ml<sup>-1</sup> phleomycin for 18 hours: no distinct RAD51 foci are detected. The other two rows are examples of similarly treated *rad51-3* <sup>-/-</sup> and *rad51-5* <sup>-/-</sup> cell lines. RAD51 was imaged by indirect immuno-fluorescence (FITC; first column); the DNA was stained with DAPI (second column); and these two were merged (third column); the cells were also visualised by differential interference contrast (DIC; fourth column) and all the images were merged (fifth column).

**Table 4-6: Percentage of RAD51 paralogue mutant cells with a specific number of RAD51 foci when treated with phleomycin.**

The percentage of cells with a specific number of sub-nuclear RAD51 foci formed (0, 1, 2, 3, 4, ≥5) when the cells were treated with 0.25 µg.ml<sup>-1</sup> or 1.0 µg.ml<sup>-1</sup> of phleomycin is shown. The different cell lines, *rad51* <sup>-/-</sup> (51<sup>-/-</sup>), *rad51-3* <sup>-/-</sup> (3<sup>-/-</sup>), *rad51-4* <sup>-/-</sup> (4<sup>-/-</sup>), *rad51-5* <sup>-/-</sup> (5<sup>-/-</sup>), and *rad51-6* <sup>-/-</sup> (6<sup>-/-</sup>) were compared with the wildtype strain 3174 (wt). The total refers to the number of cells analysed.

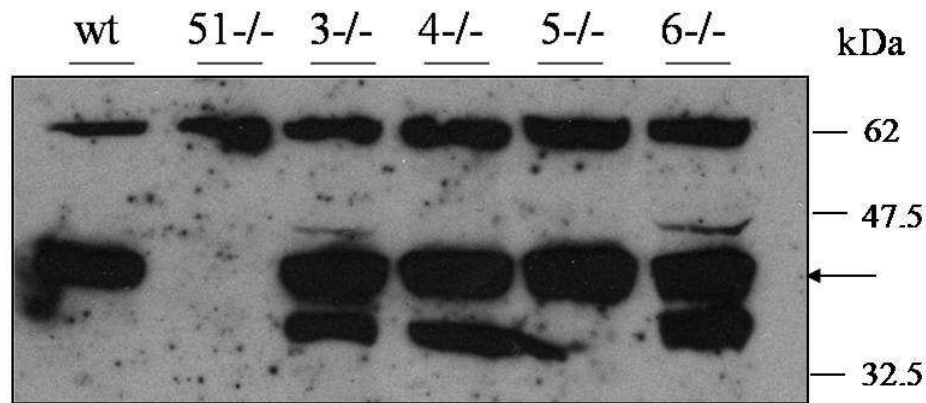
	Concentration of phleomycin	Number of RAD51 foci						Cells Counted
		0	1	2	3	4	≥5	
wt	0.25 µg.ml <sup>-1</sup>	32.0	43.0	14.0	9.0	1.0	1.0	100
	1.0 µg.ml <sup>-1</sup>	21.3	25.3	23.1	12.9	6.7	10.7	225
RAD51-3	0.25 µg.ml <sup>-1</sup>	92.9	4.1	1.0	2.0	0.0	0.0	98
	1.0 µg.ml <sup>-1</sup>	89.0	5.0	5.0	1.0	0.0	0.0	100
RAD51-4	0.25 µg.ml <sup>-1</sup>	74.3	14.4	6.9	1.7	1.3	1.3	338
	1.0 µg.ml <sup>-1</sup>	48.3	19.9	14.7	11.2	2.9	3.2	317
RAD51-5	0.25 µg.ml <sup>-1</sup>	82.2	9.9	6.9	0.0	1.0	0.0	101
	1.0 µg.ml <sup>-1</sup>	83.0	12.0	5.0	0.0	0.0	0.0	100
RAD51-6	0.25 µg.ml <sup>-1</sup>	95.9	2.2	1.0	1.0	0.0	0.0	304
	1.0 µg.ml <sup>-1</sup>	84.7	10.7	3.1	0.5	1.0	0.0	284



**Figure 4-8: Percentage of the *RAD51* paralogue mutant cells with a specific number of RAD51 foci.**

The percentage of cells with a specific number of sub-nuclear RAD51 foci formed (0, 1, 2, 3, 4, ≥5) when the cells were treated with  $1.0 \mu\text{g.ml}^{-1}$  of phleomycin is shown. The different cell lines, *rad51* -/- (51-/-), *rad51-3* -/- (3-/-), *rad51-4* -/- (4-/-), *rad51-5* -/- (5-/-), and *rad51-6* -/- (6-/-), were compared with the wildtype strain 3174 (wt).

To ensure that these results were not due to abolishment of RAD51 expression in the *RAD51* paralogue mutants, western blot analysis was performed. This showed that each of the *RAD51* paralogue mutant cell lines were expressing RAD51 (see Figure 4-9), and apparently at the same levels to each other and to the wildtype cells. Analysis to determine if RAD51 is a constitutively nuclear protein was not carried out, but one potential explanation for these findings is that lack of the RAD51 paralogues prevents RAD51 relocalisation or retention in the nucleus.



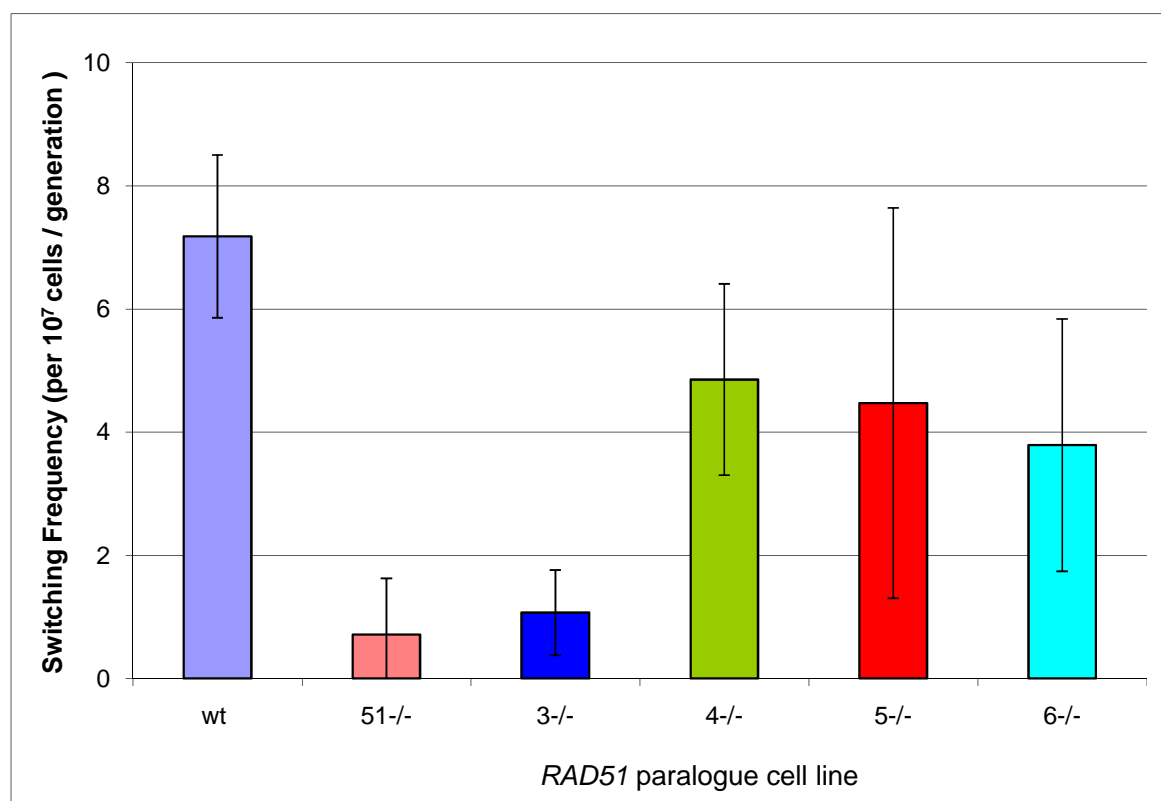
**Figure 4-9: Western blot analysis showing RAD51 expression in the *RAD51* paralogue mutant cell lines.**

Protein extracts from each cell line (wildtype (wt), *rad51*  $-/-$  (51-/-), *rad51-3*  $-/-$  (3-/-), *rad51-4*  $-/-$  (4-/-), *rad51-5*  $-/-$  (5-/-) and *rad51-6* (6-/-)) were separated on a 10 % SDS-PAGE gel and then transferred to a nylon membrane. This was then probed with affinity purified rabbit anti-RAD51 anti-serum at a dilution of 1:50 and detected using anti-rabbit anti-serum conjugated with HRP at a dilution of 1:5000 (Molecular probes). The RAD51 band is indicated with the black arrow (40.5 kDa) and the size markers are shown.

It was shown that there was a decrease in the formation of RAD51 foci in response to irradiation in *S. cerevisiae rad52* mutants, as well as mutants of both Rad51 paralogues, *rad55* and *rad57* (Gasior *et al.*, 1998). In Chinese hamster *irs1* mutant cells (*xrcc2* mutant), damage-dependent Rad51 foci formation was severely decreased compared to wildtype and complementation with a functional XRCC2 restored the Rad51 foci formation to normal (O'Regan *et al.*, 2001). This phenotype was also described in *irs1SF* mutant cells (*xrcc3* mutant), where the mutant cells did not product Rad51 foci when treated with cisplatin and  $\gamma$ -radiation (Bishop *et al.*, 1998). Studies by Takata, *et al.* confirmed these finding in the chicken cell line DT40, and showed that the Rad51 paralogue mutants, *rad51b*, *rad51c*, *rad51d*, *xrcc2* and *xrcc3*, have a decrease in the formation of RAD51 foci in response to  $\gamma$ -radiation treatment (Takata *et al.*, 2000; Takata *et al.*, 2001). In one of these studies, it was shown that mutant cells transfected with the corresponding human, *rad51c*, *xrcc2*, *xrcc3* cDNAs, or with mouse *rad51d* cDNAs were able to efficiently form Rad51 foci when treated with  $\gamma$ -radiation (Takata, *et al.*, 2001). These studies confirmed that Rad51 was present in the nucleus, and lack of focus formation was not the result of absent Rad51. These data from yeast, chicken, mammalian and *T. brucei* cells supports the proposal that the Rad51 paralogues are required for Rad51 localisation to sub-nuclear foci, or movement and formation of Rad51 on damaged DNA, or perhaps even in the stabilisation of RAD51 foci.

## 4.6 Comparison of VSG switching in *RAD51* and the *RAD51* paralogue mutants

The switching frequencies of *RAD51* paralogue homozygous mutant cell lines were analysed as described previously in McCulloch & Barry, 1999, Proudfoot & McCulloch, 2005a and Chapter 3 (Section 3.4.5 and Section 3.7.5). Due to time constraints, the VSG switching experiments were not repeated in parallel and instead the data from this study, Proudfoot & McCulloch (2005a) and McCulloch & Barry (1999) were compared for the *rad51-4* *-/-*, *rad51-6* *-/-*; *rad51-3* *-/-*, *rad51-5* *-/-*; and *rad51* *-/-* mutant cell lines respectively. In all cases, the VSG switching frequency was calculated in the same way from the raw data: isolated and cloned surviving switched *T. brucei* cells were counted over 96-well plates. This number was multiplied by 2.5, as the recovered blood used to seed the 96-well plates from the mouse was only 40 % of the total approximate blood volume of the animal. This number was then divided by the number of doubling times that had occurred during the 24 hour period following injection into the mouse and stated relative to the number of cells injected into the VSG221 immune mice. These data are shown in Figure 4-10 and Table 4-7.



**Figure 4-10: Compiled data for VSG switching frequencies of *rad51* and *RAD51* paralogue homozygous mutants.**

The graph shows the VSG switching frequencies of the different cell lines: wildtype strain 3174 (wt), *rad51* *-/-* (51-/-), *rad51-3* *-/-* (3-/-), *rad51-4* *-/-* (4-/-), *rad51-5* *-/-* (5-/-), and *rad51-6* *-/-* (6-/-). The means of triplicate data are shown, with 95 % confidence intervals indicated by the error bars.



**Table 4-7: VSG switching frequencies of *rad51* and *RAD51* paralogue homozygous mutants.**

The table shows the VSG switching frequencies of the different cell lines as the mean number of switched variants per  $10^7$  cells per generation (top row). The VSG switching frequencies are shown for wildtype strain 3174 (wt), *rad51*  $-/-$ , *rad51-3*  $-/-$ , *rad51-4*  $-/-$ , *rad51-5*  $-/-$ , and *rad51-6*  $-/-$  mutant cell line. The 95 % confidence intervals are indicated by  $\pm$ .

wt	<i>rad51</i> $-/-$	<i>rad51-3</i> $-/-$	<i>rad51-4</i> $-/-$	<i>rad51-5</i> $-/-$	<i>rad51-6</i> $-/-$
7.18 ( $\pm$ 1.32)	0.44 ( $\pm$ 0.33)	1.07 ( $\pm$ 0.69)	4.85 ( $\pm$ 1.55)	4.47 ( $\pm$ 3.17)	3.79 ( $\pm$ 2.05)

This analysis suggests firstly, that all the mutants have to some extent, an impaired ability to undergo VSG switching relative to wildtype cells. Secondly, the data suggests that *rad51*  $-/-$  and *rad51-3*  $-/-$  mutants have a lower mean VSG switching frequency compared with the three other *RAD51* homozygous mutants, *rad51-4*  $-/-$ , *rad51-5*  $-/-$  and *rad51-6*  $-/-$  (see Table 4-7). In the cases of *rad51-3*  $-/-$ , *rad51-4*  $-/-$  and *rad51-6*  $-/-$ , these impairments are also seen relative to the heterozygous mutants from which the homozygous mutants were derived (Chapter 3, Section 3.4.5 and 3.7.5, Proudfoot & McCulloch, 2005a). For *rad51-5*  $-/-$ , this was not the case, and it was concluded that *RAD51-5* may not act in VSG switching (Proudfoot & McCulloch, 2005a). This data also suggests that *RAD51-4* and *RAD51-6* may not act in VSG switching.

The 95 % confidence intervals indicated by the error bars are large for the *rad51-4*  $-/-$ , *rad51-5*  $-/-$  and *rad51-6*  $-/-$  cell lines. This is most likely due to the high variability and nature of the *in vivo* experiment. Nevertheless, one interpretation is that a distinct phenotype is seen, whereby loss of *RAD51* or *RAD51-3* substantially impairs the ability of VSG switching although some switching events still occurred. In contrast, loss of *RAD51-4*, *RAD51-5* and *RAD51-6* has less of an impact. The *rad51*  $-/-$  and *rad51-3*  $-/-$  phenotypes are very similar to those described for the *brca2*  $-/-$  mutant (Hartley & McCulloch, 2008), where there was a strong and comparable reduction in VSG switching. This suggests that there are a considerable number of factors that regulate *RAD51* function during antigenic variation which may suggest that it is a randomly initiated process, where many forms of initiation can induce VSG switching. This may explain why all the mutants examined can still undergo antigenic variation by expression site and VSG gene conversion reactions, mechanisms which require HR. It rests on the assumption, however, that there are multiple roles for the *RAD51* paralogues in recombination, and these are reflected in antigenic variation. This assumption requires further testing.

## 4.7 Summary

The aim of this summary chapter was to show direct comparisons between mutant cell lines of all four RAD51 paralogues, RAD51 and wildtype cell lines. With one exception, all these experiments were carried out at the same time, allowing direct comparisons. The data described in this chapter confirms all previously published data by McCulloch & Barry (1999) and Proudfoot & McCulloch (2005a), as well as the results described in Chapter 3 of this thesis. Only small differences in quantification of specific phenotypes were seen.

The results described here verify that all four *T. brucei* RAD51 paralogues have a role in DNA damage repair, with the mutants having an increased sensitivity to DNA damaging agents. Most likely, this is due to impairment in the parasite cell's ability to undergo HR, as this pathway was found to be operating at reduced efficiency in each mutant. It is also likely that the decreased growth rate of the mutant cells stems from the above repair/recombination defect. It has also been confirmed that the *T. brucei* RAD51 paralogues have a part to play in the formation or stabilisation of phleomycin induced RAD51 foci formation. Although the phenotypes of all the *RAD51* paralogue mutants are similar to one another, they are not equivalent. Two of the RAD51 paralogue proteins, RAD51-4 and RAD51-5, appear have different functions within the DNA repair pathways. The growth rates of *rad51-4* *-/-* and *rad51-5* *-/-* mutant cells are quicker than that of the *rad51* and the other *RAD51* paralogue homozygous mutants. Furthermore, *rad51-4* *-/-* appears to be more resistant to the DNA-damaging agent, phleomycin, and is more able to produce DNA-damage induced RAD51 foci than the other RAD51 paralogue mutants. The implications of these results are unknown, but may suggest that they have a role or roles which can be performed by some other protein, such as another RAD51 paralogue. It may also be the case that the role of RAD51-4, is aiding or stabilising the functions of other RAD51 paralogues see Chapter 5 (Section 5.8).

These suggestions are only speculative, and there is the need to develop these theories. It will be important to examine the direct roles of each RAD51 paralogue by biochemical analysis to build a better picture regarding their function, such as testing the DNA binding ability or ATP hydrolysis capacity of the proteins, as well as assessing how they influence RAD51 function *in vivo*. Other potentially complementary experiments could be carried out directly in *T. brucei*: over-expression of RAD51 in the RAD51 paralogue mutant cells would determine if they have indispensable roles in aiding RAD51 function; the generation of double knockout mutants of the RAD51 paralogues would distinguish between potential

functions in one pathway, or their action in separate pathways. These experiments would enable a complete model of the roles of the *T. brucei* RAD51 paralogues to be formed.

**Chapter 5. Examining the  
associations between *T. brucei*  
RAD51 and the RAD51  
paralogues**

## 5.1 Introduction

In *T. brucei*, there are four RAD51 paralogue proteins, RAD51-3, RAD51-4, RAD51-5, and RAD51-6, which share limited sequence homology with each other and RAD51 and DMC1 (Proudfoot & McCulloch, 2005a). To date, no analysis has been done to evaluate if and how these *T. brucei* RAD51 paralogues interact, although there is good evidence to suggest that each RAD51 paralogue in *T. brucei* has a role in DNA damage repair requiring the homologous recombination (HR) pathway (see Chapter 3, Chapter 4 and Proudfoot & McCulloch, 2005a). Extensive study of the mammalian Rad51 paralogues (XRCC2, XRCC3, Rad51B, Rad51C and Rad51D) confirms that they physically interact with each other, possibly providing discrete protein complexes that act in DNA recombination (Liu *et al.*, 2002b; Schild *et al.*, 2000). The current model for mammalian Rad51 paralogue interactions suggests two main complexes, incorporating all five paralogues, with Rad51C interacting with XRCC3 (CX3) in one complex and Rad51B-Rad51C-Rad51D-XRCC2 (BCDX2) in another (Liu *et al.*, 2002b; Masson *et al.*, 2001a). Rad51C appears to play an important role, and is a component of both complexes (Liu *et al.*, 2002b; Miller *et al.*, 2004). The extent of the conservation of these two complexes in other eukaryotes has not been examined, but preliminary results show that two of the *Arabidopsis thaliana* Rad51 paralogues interact: AtRad51C and AtXRCC3. This study also showed that AtXRCC3 interacts with AtRad51 (Osakabe *et al.*, 2002). In *Saccharomyces cerevisiae*, the two RAD51 paralogues, RAD55 and RAD57, form a single stable complex (Hays *et al.*, 1995; Johnson & Symington, 1995). By these precedents, it is likely that the *T. brucei* proteins will also form a complex or complexes. However, as there are low levels of sequence homology among these proteins between the species, the interactions are not simple to predict. In addition, the potential conservation of such complexes in other eukaryotes, such as *Drosophila*, has not been examined (Abdu *et al.*, 2003; Ghabrial *et al.*, 1998).

To determine if the *T. brucei* RAD51 paralogues interact with each other and to establish the number and composition of such complexes, three approaches were adopted. The first was yeast two-hybrid analysis, where interaction between RAD51 and the RAD51 paralogues was assayed using two reporter systems,  $\beta$ -galactosidase activity and histidine auxotrophy. Second, co-immunoprecipitation was employed; antisera were raised against His-tagged recombinant RAD51 and RAD51 paralogue proteins. These antisera were then used in co-immunoprecipitation analyses to determine direct *in-vivo* interaction. Third, by generating *T. brucei* strains expressing epitope-tagged RAD51 paralogues, *in-vivo* co-purification analyses were performed.

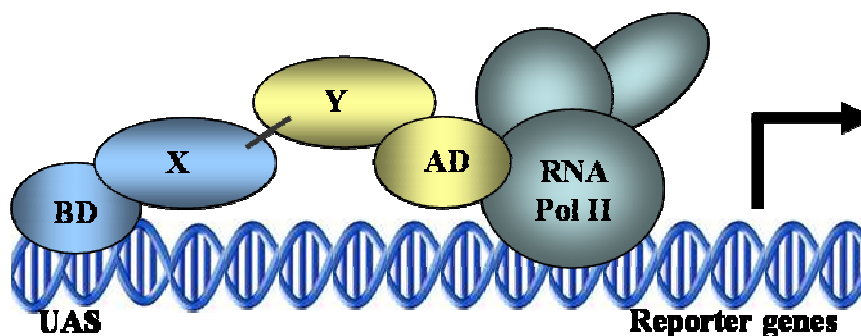
## 5.2 Principles of the yeast two-hybrid system

The yeast two-hybrid system is used to detect protein-protein interactions, exploiting the knowledge that most eukaryotic transcription activator proteins have two functional domains, a DNA-binding domain and an activating domain (Fields & Song, 1989). The DNA-binding domain attaches to a DNA sequence called the upstream activation site (UAS) and the activating domain recruits the RNA polymerase II machinery to allow expression of the genes downstream of the UAS (Luban & Goff, 1995; Figure 5-1). Individually these domains cannot activate gene expression, and both domains must be in close proximity to the DNA for activation. GAL4 is a budding yeast transcriptional activator with a DNA-binding and activating domain, and is required for the activation of the *GAL* genes in response to galactose (Giniger *et al.*, 1985). It has been proposed that the first 147 amino acid residues of GAL4 is the DNA-binding domain and the last 113 residues is the activating domain, showing that up to 80 % of GAL4 protein could be deleted without drastic loss of transcriptional activation function (Ma & Ptashne, 1987a).

It has also been shown that transcriptional activation can occur when the domains originate from independent and different polypeptides, and prokaryotic proteins can provide the function for either or both of the DNA-binding and activating domains. For instance, Brent & Ptashne (1985) created a functional transcription activator protein in yeast from the DNA-binding domain from the *Escherichia coli* repressor protein LexA fused to the activating domain of *S. cerevisiae* GAL4. The same group found that yeast transcriptional activators could be encoded by non-homologous DNA fragments from *E. coli*, fused to the encoding sequence of the DNA-binding domain of yeast GAL4 (Brent & Ptashne, 1985; Ma & Ptashne, 1987b).

Fields & Song (1987) first suggested constructing independent fusions of the two potentially interacting proteins, with one protein fused to the DNA-binding domain, and the other protein to the activating domain. In this study, the authors used two proteins known to interact with each other, SNF1 and SNF4, the former fused to the DNA-binding domain of the GAL4 protein amino acids (1-147) and the latter fused to the activating domain of GAL4 (768-881) (In Figure 5-1, SNF1 would be X and SNF4 would be Y). The two vectors expressing fusion polypeptides were introduced into a yeast strain that had the native *GAL4* gene deleted and also contained a *GAL1-LacZ* reporter gene in the yeast genome. The SNF1 fusion to the DNA-binding domain was defined as the “bait” protein and the SNF4 fusion to the activating domain was defined as the “prey” protein. As SNF1

and SNF4 interact, the activating domain came into close proximity with the DNA-binding domain, and activated transcription of the reporter gene, *LacZ* encoding  $\beta$ -galactosidase.



**Figure 5-1: Model of the yeast two-hybrid system.**

BD: DNA binding domain LexA, X: Bait protein, AD: Activation domain B42, Y: Prey protein, RNA pol II: Yeast RNA polymerase II and associated proteins, UAS: Upstream activation sequence. The black arrow indicates transcription of the reporter genes. (Adapted from: Causier & Davies, 2002).

These initial experiments demonstrated the possibility of using GAL4 to examine protein-protein interaction and led to the discovery of other DNA-binding and activating domains (Fields & Song, 1989). Indeed, yeast two-hybrid analysis is now widely used to study protein interaction networks in many organisms (Gyuris *et al.*, 1993; Vojtek *et al.*, 1993). The work in this thesis used LexA as the DNA-binding domain (BD) and B42 as the activating domain (AD), (Figure 5-1).

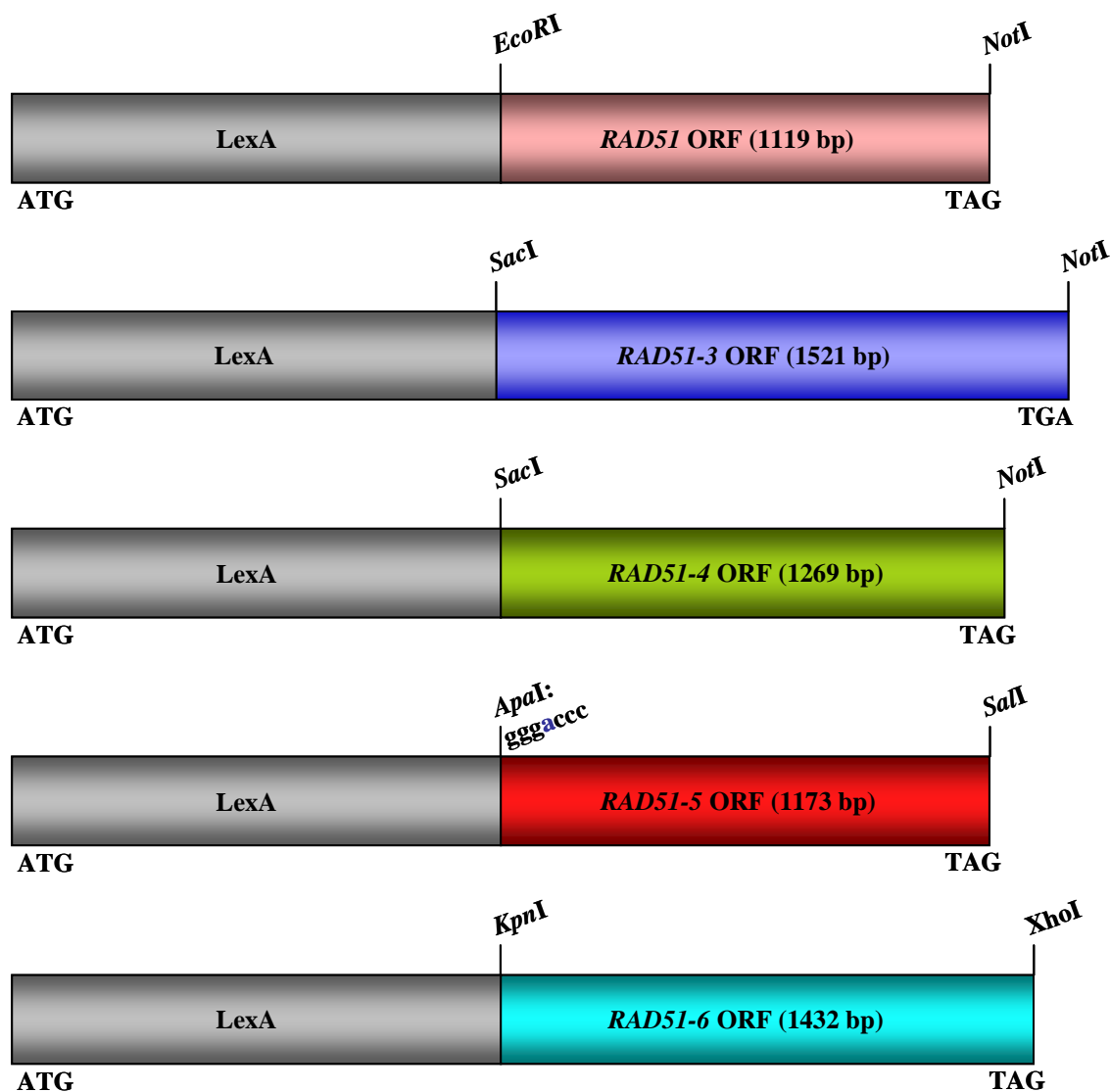
## 5.3 Components of the yeast two-hybrid system

### 5.3.1 The role of LexA

The LexA protein is a bacterial repressor protein which binds to DNA (Brent & Ptashne, 1980). This DNA binding protein acts as a dimer, and recognises the LexA operator sequence. Brent & Ptashne (1985) showed that a fusion of a fragment of the *E. coli* LexA protein (1-87 amino acids) to the *S. cerevisiae* Gal4 activating domain (74-881 amino acids) activated transcription of the reporter genes. In contrast, LexA alone did not activate transcription when the LexA binding sites were cloned upstream of the reporter genes in the yeast genome. LexA enters the yeast nucleus independently and further studies have validated that *E. coli* LexA could be used as a DNA-binding domain in the application of yeast two-hybrid screens (Vojtek *et al.*, 1993).

Invitrogen's yeast two-hybrid system vector, pHybLex/Zeo (see Materials and Methods, Section 2.10) was used to provide the DNA-binding domain. In this, the full length LexA protein (202 amino acids) was expressed, and the *Streptoalloteichus hindustanus ble* gene under the control of *TEF1* and *EM-7* promoters, allowed zeocin resistance selection in *S.*

*cerevisiae* and *E. coli*, respectively. Using this vector to provide the DNA-binding domains, LexA fusions were made with *T. brucei* RAD51 and the four *T. brucei* RAD51 paralogues. Figure 5-2 shows the fragment size of each ORF and also shows the enzyme restriction sites used in the cloning of each ORF into pHybLex/Zeo. pHybLex/Zeo-RAD51, pHybLex/Zeo-RAD51-3, pHybLex/Zeo-RAD51-4, pHybLex/Zeo-RAD51-6 were made by Sandra Terry. These constructs were confirmed to be correct and to be in frame with LexA by sequence analysis, with the exception of LexA-RAD51-5 (which is described in Section 5.4.4).



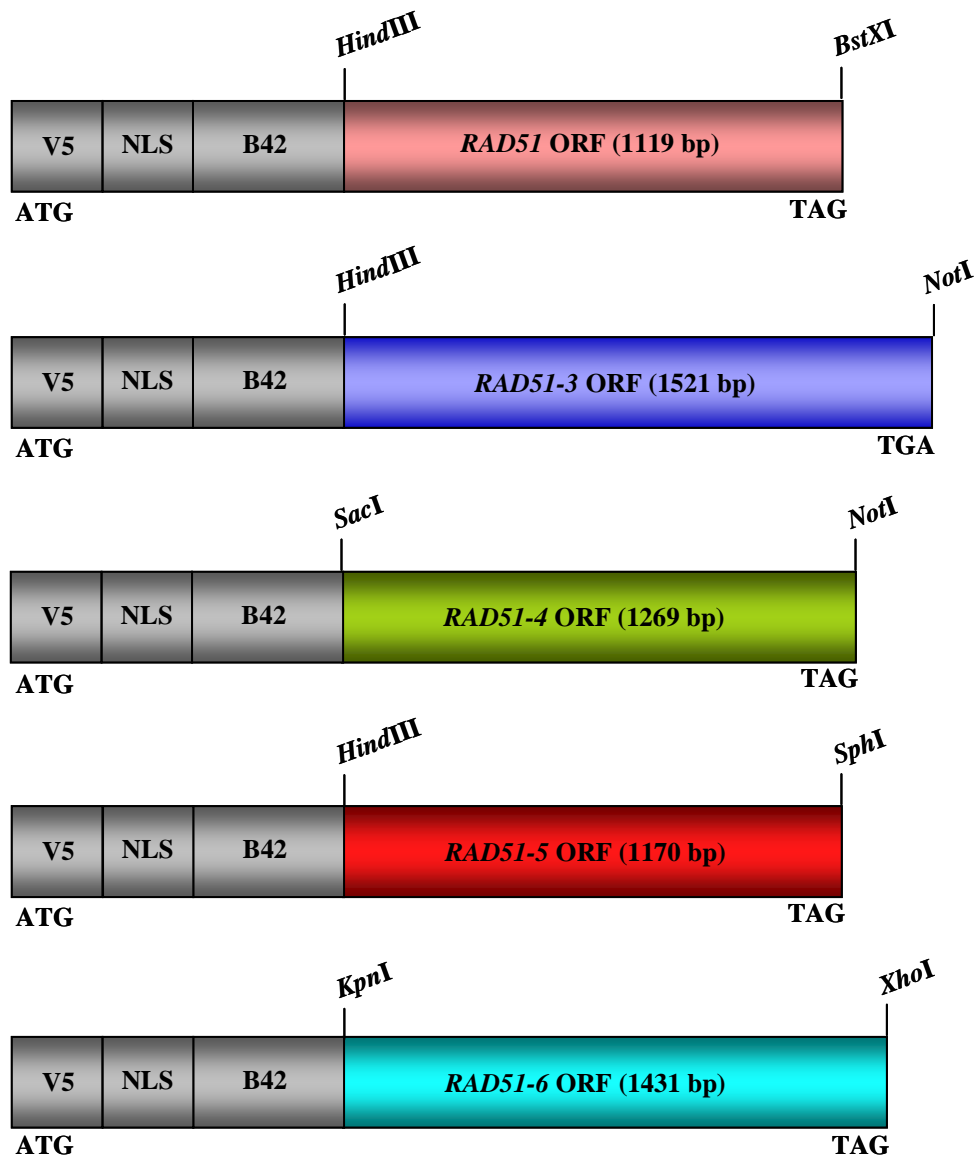
**Figure 5-2: Diagram showing the LexA fusions with RAD51 and RAD51 paralogues.**

The complete *lexA* gene (606 bp) is part of the pHybLex/Zeo vector and allows expression of a fusion protein with the LexA DNA-binding domain. The LexA polypeptide adds 32 kDa to the N terminus of RAD51 and the RAD51 paralogue proteins in this configuration. The length of the ORF of RAD51 and each RAD51 paralogue, and the restriction sites used for the cloning into the vector pHybLex/Zeo, are shown.



### 5.3.2 The role of B42

Ma & Ptashne (1987b) identified a number of different activating domains when the DNA-binding domain of the GAL4 gene was fused to polypeptides encoded from random DNA fragments of the *E. coli* genome. All transcriptional activating sequences identified were acidic, having a net negative charge (Ma & Ptashne, 1987b). One of these random DNA fragments, named B42, was shown to be amongst the strongest activator polypeptides when fused to the DNA-binding domain of GAL4 (Ma & Ptashne, 1987b). A fusion of the DNA binding domain of LexA and B42 was also generated, and was shown to activate a *lacZ* reporter gene downstream of a *lexA* operator. This confirmed that DNA-binding and activator polypeptides from *E. coli* could be used to activate genes on the yeast chromosome (Ma & Ptashne, 1987b), with the advantage that they should not be subjected to regulation by other yeast proteins. A further study confirmed that B42 could be used as an activating domain in the yeast two-hybrid analysis (Gyuris *et al.*, 1993). In the Invitrogen system used here, the pYesTrp2 vector expresses a fusion of the haemagglutinin epitope tag V5, the SV40 nuclear-localisation sequence (NLS) and B42 as the activating domain. The complete ORF of *T. brucei RAD51* and the *RAD51* paralogues were cloned downstream of, and in frame with, the V5-NLS-B42 ORF to generate N terminal fusions with the three polypeptides (see Figure 5-3), this was performed by Sandra Terry. Also present in pYesTrp2 is the *TRP1* gene, which allows for auxotrophic selection of the vector in a tryptophan negative yeast host, and the ampicillin resistance gene which allows selection in *E. coli*.



**Figure 5-3: Diagram showing the V5-NLS-B42 fusions with RAD51 and the RAD51 paralogs.**

The V5 epitope allows detection of the fusion proteins using Anti-V5 antibody. SV40 large T antigen nuclear localisation sequence (NLS) allows the V5-NLS-B42-RAD51 fusion and V5-NLS-B42-RAD51 paralogue fusions to enter the nucleus for potential interactions with LexA-RAD51 and LexA-RAD51 paralogue fusions. B42 is the transcriptional activation domain. The V5-NLS-B42 peptide adds 12 kDa to the RAD51 and the RAD51 paralogue polypeptides. The length of the ORF of *RAD51* and each *RAD51* paralogue, and the restriction sites used for cloning into the vector pYesTrp2 are shown.

### 5.3.3 *S. cerevisiae* L40 strain

The *S. cerevisiae* L40 strain is commonly used with LexA-based yeast two-hybrid systems (Invitrogen). The strain contains two reporters, *HIS3* and *lacZ*, both under the control of a minimal *GAL1* promoter. Both reporter genes have multiple copies of the *lexA* operator sequences that act as upstream activating sequences (UAS): four copies are positioned upstream of the *HIS3* reporter gene and eight copies are upstream of the *lacZ* reporter gene (Vojtek *et al.*, 1993). The L40 strain is auxotrophic for histidine (his), adenine (ade), tryptophan (trp), and prototrophic for leucine (leu) and uracil (ura). The requirement for tryptophan allows selection of transformants containing the pYesTrp2 vector.

### 5.3.4 Independent testing of protein interactions

Protein interaction of the RAD51 and RAD51 paralogues was examined by three independent tests, histidine auxotrophy, and two methods of testing  $\beta$ -galactosidase activity: filter lift assay and Pierce  $\beta$ -galactosidase kit. Invitrogens pHybLex/Zeo-Fos2 and pYesTrp-Jun are positive controls and interact in the yeast-two hybrid system, while pHybLex/Zeo-Lamin is the negative control for the bait plasmid. These positive and negative controls for the yeast-two hybrid analysis were not done. Nevertheless, the three independent methods of interaction analyses were performed and two independent transformants for each test were carried out. As a negative control the empty vector pYesTrp2 was used in combination with the pHybLex vectors. To verify that all fused proteins were expressed western analyses were performed on the cell lysates co-transformed with the LexA and V5 fusions. In addition, the histidine auxotrophy tests were performed at least twice.

#### 5.3.4.1 To test for histidine auxotrophy

One method to examine protein-protein interactions in the yeast two-hybrid system used here involves testing whether or not different co-transformants of pHybLex/Zeo and pYesTrp2 fusion vectors grow in the absence of histidine. As mentioned above, the L40 strain used is auxotrophic for histidine and contains the *HIS3* reporter gene under the control of minimal *GALI* promoters fused to *lexA*-operator sequences. If the LexA and V5-NLS-B42 fused proteins interact, transcription of the minimal *GALI* promoter allows expression of imidazole-glycerolphosphate dehydratase (HIS3), which catalyses the sixth step in histidine biosynthesis (Alifano *et al.*, 1996; SGD-project, *Saccharomyces* Genome Database), allowing growth on media lacking histidine. Use of a “minimal” *GALI* promoter allows identification of protein-protein interactions that may otherwise be toxic to yeast growth (Invitrogen manual). In the following experiments to test for histidine auxotrophy, 3-amino-1,2,4-triazole (3' aminotriazole) was added to the histidine negative (his-) selective media plates (see Materials and Methods, Section 2.10.1). 3' aminotriazole is a competitive inhibitor of imidazole-glycerolphosphate dehydratase, the product of the *HIS3* gene (Kanazawa 1998), and allows titration of the level of *HIS3* expression. Non-specific and very weak interactions will produce some imidazole-glycerolphosphate dehydratase, which would result in false positives in the test. To circumvent this, small and increasing concentrations (5 mM, 10 mM, and 15 mM) of 3' aminotriazole were added. Specific and strong interactions would produce more imidazole-glycerolphosphate dehydratase and as a result, co-transformants expressing ‘true’ interactors will grow on the his- media plates even in the presence of the highest concentration of 3' aminotriazole.

### **5.3.4.2 To test for $\beta$ -galactosidase activity using the filter lift assay**

Since the *S. cerevisiae* L40 strain also contains a *lacZ* marker gene under the control of a minimal *GAL1* promoter fused to eight copies of the LexA-operator sequences, interaction between and amongst RAD51 and the RAD51-paralogues can be assayed through induction of  $\beta$ -galactosidase activity. This was assayed in two ways: the  $\beta$ -galactosidase filter lift assay and the yeast  $\beta$ -galactosidase assay kit (Pierce).

In the first method,  $\beta$ -galactosidase activity was determined qualitatively using the X-gal filter lift assay (Breedon & Nasmyth, 1985). This is a quick screening method based on a blue/white visualisation of yeast colonies immobilized on filters (see Materials and Methods, Section 2.10.2). Expression of the *lacZ* should only occur if RAD51 and the RAD51 paralogue fused proteins interact, resulting in blue colouration of the colonies. This method uses 5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactoside commonly known as X-Gal. Colourless X-gal is hydrolyzed by the  $\beta$ -galactosidase enzyme, forming 5-bromo-4-chloro-3-hydroxyindole. This by-product is then oxidized, resulting in an insoluble blue product, 5,5'-dibromo-4,4'-dichloro-indigo. This method also allows imaging of the plates, and due to its sensitivity enables detection of low levels of  $\beta$ -galactosidase activity (Lehninger *et al.*, 1993).

### **5.3.4.3 To test for $\beta$ -galactosidase activity using a “Yeast $\beta$ -galactosidase assay kit”**

In the second method, the  $\beta$ -galactosidase activity was determined quantitatively using a yeast  $\beta$ -galactosidase assay kit from Pierce. The kit was used to detect and quantify transcriptional activation of the  $\beta$ -galactosidase gene to examine the extent of any interactions between RAD51 and the RAD51 paralogue fused proteins. The same kit was used for all the analyses, and all required solutions were provided and used according to manufacturer's instructions.  $\beta$ -galactosidase activity was assayed by measuring the cleavage of the lactose like compound ONPG (*o*-nitrophenyl- $\beta$ -D-galactopyranoside) by the  $\beta$ -galactosidase enzyme to produce *o*-nitrophenol. *o*-nitrophenol is an insoluble yellow precipitate and its generation was quantified by spectrophotometry at an absorbance of 420 nm. The measurement of  $\beta$ -galactosidase activity considers the optical density of the cells, the amount of cells used, and time of incubation of the reactions using the following equation:

$$(1000 \times \text{Absorbance } 420 \text{ nm}) \div (t \times V \times \text{OD } 620\text{nm})$$

where  $t$  is the assay reaction time (in minutes) and  $V$  is the volume of culture used in the assay (in ml). The enzymatic reaction time varies depending on the levels of the  $\beta$ -galactosidase expression. The cell volume and density will also affect the  $\beta$ -galactosidase activity units. This information allows quantification of the  $\beta$ -galactosidase activity within each co-transformant, and comparisons are easily carried out.

## 5.4 Results for yeast two-hybrid assays

To examine the interactions between all the combinations of *T. brucei* RAD51 and the four RAD51 paralogues, each corresponding ORF was cloned into the two yeast two-hybrid vectors (pHybLex/Zeo and pYesTrp2; Invitrogen) which were described in Section 5.3.1 and Section 5.3.2. This should allow expression of each full length protein as an N terminal fusion with the LexA-DNA binding domain from pHybLex/Zeo, or with the V5-NLS-B42 activating domain from pYesTrp2. All pair-wise combinations of pHybLex/Zeo-RAD51 and the pHybLex/Zeo-RAD51 paralogue vectors were then co-transformed into *S. cerevisiae* L40 cells with the pYesTrp2-RAD51 and pYesTrp2-RAD51 paralogue vectors (Materials and Methods, Section 2.4.6). Two independent co-transformants of each pair-wise combination were analysed, and named X and Y. As a negative control, each of the pHybLex/Zeo-RAD51 and the pHybLex/Zeo-RAD51 paralogue vectors were co-transformed with the empty activator vector pYesTrp2 (EV). If V5-NLS-B42-RAD51 or V5-NLS-B42-RAD51-paralogue fusions act as interaction-dependent activating domains the reporter genes are transcribed, i.e. the reporter genes should remain silent if there is no interaction, unless the LexA fusions act as activators themselves.

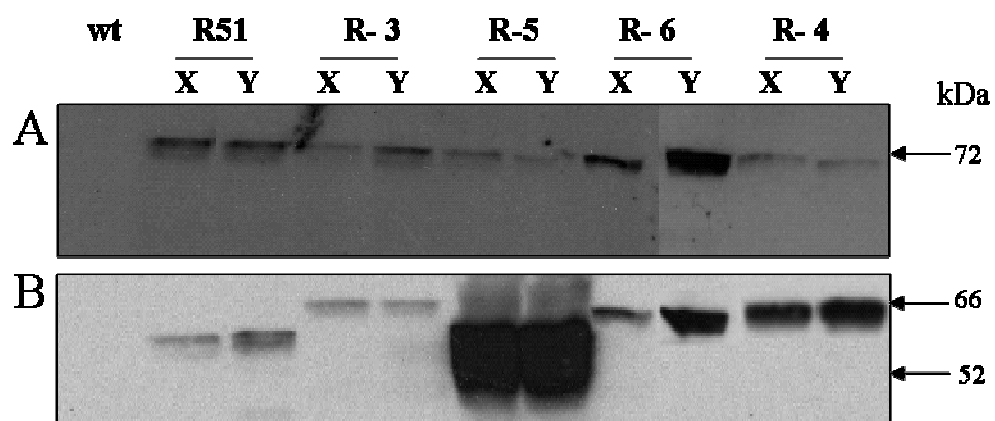
Following selection for the presence of the two vectors, western blots were performed to test for expression of both LexA and V5-NLS-B42 fusions in the transformants prior to examination of their interaction by histidine auxotrophy and  $\beta$ -galactosidase activity.

### 5.4.1 Yeast two-hybrid analysis with LexA-RAD51 as the DNA-binding domain fusion

To begin examining potential interactions, RAD51 was first used as the “bait” (expressed as a LexA fusion) and its interactions assayed with RAD51 and the RAD51 paralogues as the “prey” (expressed as V5-NLS-B42 fusions). To do this, pHybLex/Zeo-RAD51 was co-

transformed with pYesTrp2-RAD51, pYesTrp2-RAD51-3, pYesTrp2-RAD51-4, pYesTrp2-RAD51-5 and pYesTrp2-RAD51-6; see Materials and Methods, Section 2.4.6. As a negative control, pHybLex/Zeo-RAD51 was also co-transformed with empty vector pYesTrp2. Two independent co-transformants (X and Y) were analysed for each combination to determine if the LexA-RAD51 fusion protein was expressed as well as the V5-NLS-B42-RAD51 fusion and the V5-NLS-B42-RAD51 paralogue fusion proteins, (see Figure 5-4). To do this, cell extracts were prepared from each of the co-transformed cells, and the proteins were separated on a 10 % SDS-PAGE gel. The proteins were then transferred to a nylon membrane and probed with goat anti-Lex IgG anti-serum (Santa Cruz) at a 1:1000 dilution and detected with anti-goat IgG-HRP anti-serum (Santa Cruz) at a 1:5000 dilution (Figure 5-4A). The blot was then stripped and reprobed with mouse anti-V5 IgG anti-serum (Invitrogen) at a dilution of 1:5000 and detected with anti-mouse IgG-HRP anti-serum (Sigma) at a 1:2500 dilution (Figure 5-4B).

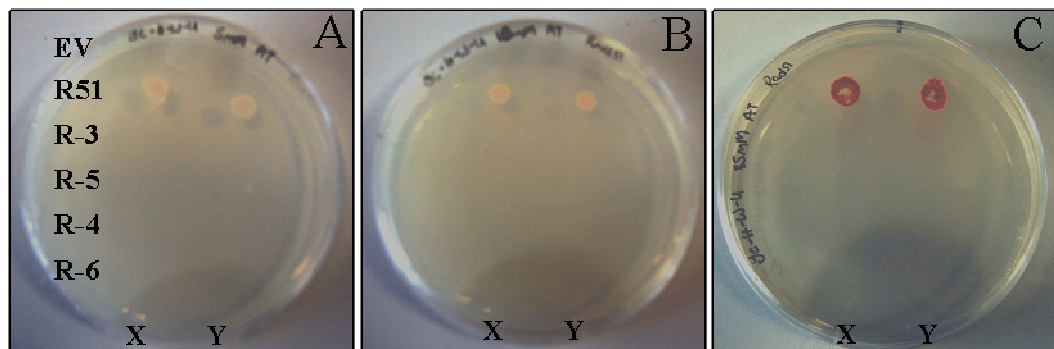
In all the co-transformants, a protein of the expected size for the LexA-RAD51 fusion was detected (72 kDa) by the anti-LexA anti-serum (Figure 5-4). This LexA-RAD51 fusion was absent in the untransformed *S. cerevisiae* L40 wildtype cell extract. Proteins of the expected sizes of RAD51 and the four RAD51 paralogs expressed as V5-NLS-B42 fusions were also detected by the anti-V5 anti-serum. It is unclear why the apparent expression or reactivity of the V5-NLS-B42-RAD51-5 fusion is much greater than the others, but it must be pointed out that no loading controls were performed.



**Figure 5-4: Western analysis of LexA-RAD51 expression with V5-NLS-B42-RAD51 and the RAD51 paralogs in *S. cerevisiae* L40 co-transformants.**

Two co-transformants (X and Y) expressing LexA-RAD51 and each V5-NLS-B42 fusion of RAD51 (R51), RAD51-3 (R-3), RAD51-5 (R-5), RAD51-6 (R-6) or RAD51-4 (R-4) are shown. **Figure 5-4A:** Protein extracts from the co-transformants, and from the untransformed *S. cerevisiae* L40 cells (wt) were separated by SDS-PAGE, blotted and probed with anti-LexA anti-serum. **Figure 5-4B:** The blot was then stripped and reprobed with anti-V5 anti-serum. The inferred sizes of the proteins (kDa) are between 52 and 66 kDa as shown. The predicted sizes of the fusion polypeptides are as follows: LexA-RAD51: 72 kDa and V5-NLS-B42 fusions- RAD51: 52 kDa, RAD51-3: 66.5 kDa, RAD51-4: 57.7 kDa, RAD51-5: 53.6 kDa and RAD51-6: 62.4 kDa. The fusion of V5- NLS-B42 adds 12 kDa and the fusion of LexA adds 32 kDa to the proteins.

To examine interactions, the histidine auxotrophy test was first used and showed that RAD51 interacts with itself as revealed by the clear growth on the plates lacking histidine and containing 15 mM 3' aminotriazole (Figure 5-5). In contrast, there was no evidence that RAD51 interacts with any of the other RAD51 paralogues in this system, as there is an absence of growth in all these co-transformants for reasons discussed in Section 5.3.4.1. To ensure that the cultures of each co-transformant were viable, a positive control plate was used. The cultures were plated on media containing all the essential amino acids, including histidine. As histidine was provided, all the inoculated cultures grew as small circular colonies. This plate was then used in the X-gal filter lift assay, see Figure 5-6.



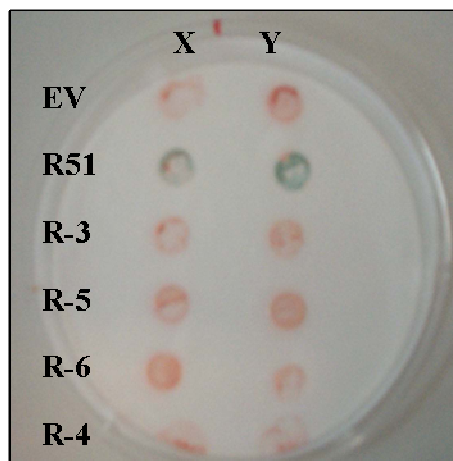
**Figure 5-5: Histidine auxotrophy analysis of LexA-RAD51 with V5-NLS-B42-RAD51 and the V5-NLS-B42-RAD51 paralogues.**

**Figure 5-5** shows growth or absence of growth of the LexA-RAD51 fusion with V5-NLS-B42-RAD51 and each V5-NLS-B42-RAD51-paralogue fusion on medium lacking histidine. The two co-transformants (X and Y) are shown for LexA-RAD51 with each combination of V5-NLS-B42 fusion [Labelled: RAD51 (R51), RAD51-3 (R-3), RAD51-5 (R-5), RAD51-6 (R-6) or RAD51-4 (R-4)]. As a negative control, the LexA-RAD51 expression vector was co-transformed with the pYesTrp2 vector (empty vector: EV). The plates have three different concentrations of histidine inhibitor, 3' aminotriazole, **Figure 5-5A**: 5 mM, **B**: 10 mM, **C**: 15 mM.

The histidine auxotrophy test was performed at least twice and the plates in Figure 5-5 were part of multiple experiments. It is noteworthy that the L40 strain has a mutation in *ADE2* gene and as a result cannot produce phosphoribosylaminoimidazole carboxylase. This carboxylase is a requirement of the adenine biosynthetic pathway converting 5'-phosphoribosyl-5-aminoimidazole to 5'-phosphoribosyl-5-aminoimidazole-4-carboxylate (Rolfes, 2006; Ugolini & Bruschi, 1996). As a result, the L40 strain is auxotrophic for adenine, as mentioned in Section 5.3.3 and addition of adenine to the medium is required for growth of the wildtype and co-transformants. An accumulation of 5'-phosphoribosyl-5-aminoimidazole in the yeast cell results in a pink/red pigment, which is clearly seen in Figure 5-5C.

In order to verify the results of the histidine auxotrophy test,  $\beta$ -galactosidase activity was analysed. As a positive control for the histidine auxotrophy test, the cultures of each co-

transformant were plated on media containing all the essential amino acids, including histidine. These colonies were then analysed using the X-gal filter lift assay (see Material and Methods, Section 2.10.2). In the transformant expressing LexA-RAD51 and V5-NLS-B42-RAD51, there was a clear colour change from pink to blue in the X-gal filter lift assay, confirming self-interaction of RAD51 (Figure 5-6). No such evidence for  $\beta$ -galactosidase activity was seen in any of the other transformants, or in the control cells expressing only LexA-RAD51. These results are consistent with the histidine auxotrophy assay.

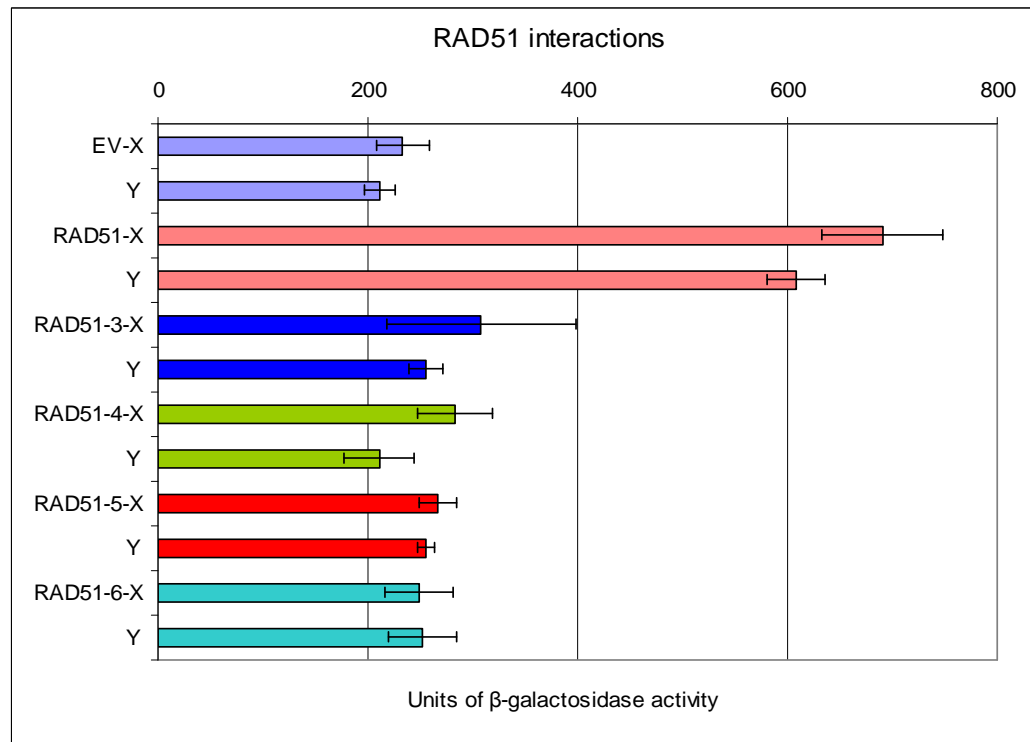


**Figure 5-6: X-gal filter lift analysis of LexA-RAD51 with V5-NLS-B42-RAD51 and the V5-NLS-B42-RAD51 paralogues.**

As a positive control for the histidine auxotrophy test, the cultures of each co-transformant (X and Y) were plated on medium containing all the essential amino acids. These transformants were then analysed using the X-gal filter lift assay. LexA-RAD51 was co-expressed with each combination of V5-NLS-B42 fusion [Labelled: RAD51 (R51), RAD51-3 (R-3), RAD51-5 (R-5), RAD51-6 (R-6) or RAD51-4 (R-4)]. As a negative control, the LexA-RAD51 expression vector was co-transformed with the pYesTrp2 vector (empty vector: EV). Interactions are identified by a colour change from pale pink to blue, where  $\beta$ -galactosidase expression was induced from a *lacZ* reporter gene.

Finally, the interaction of LexA-RAD51 with the V5-NLS-B42 RAD51 and V5-NLS-B42 RAD51 paralogue fusions were quantified by measuring the  $\beta$ -galactosidase activity in the lysate of the co-transformants, using the yeast  $\beta$ -galactosidase assay kit by Pierce (see Materials and Methods, Section 2.10.3). As a negative control, the LexA-RAD51 expression vector was co-transformed with the pYesTrp2 vector (empty vector: EV). The results of this are shown in Figure 5-7 and again demonstrate RAD51 self interaction, and no other interaction between RAD51 and any of the RAD51 paralogues. Only in the two co-transformants (X and Y) expressing LexA-RAD51 and V5-NLS-B42-RAD51 were the levels of the  $\beta$ -galactosidase higher than the control expressing only LexA-RAD51.





**Figure 5-7: Quantification of β-galactosidase activity of LexA-RAD51 with V5-NLS-B42-RAD51 and the V5-NLS-B42-RAD51 paralogues.**

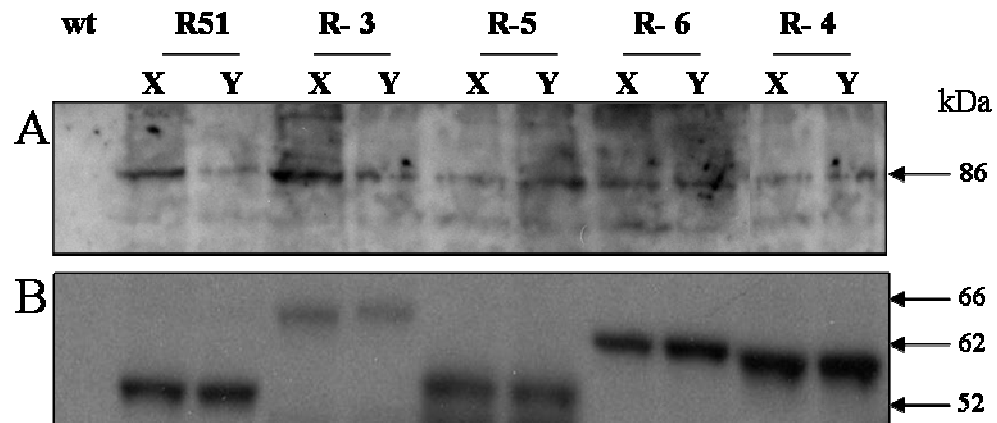
β-galactosidase activity was analysed for cells expressing the LexA-RAD51 fusion with V5-NLS-B42-RAD51 and each V5-NLS-B42-RAD51-paralogue fusion using the yeast β-galactosidase assay kit by Pierce. The measurement of β-galactosidase activity considers the generation of *o*-nitrophenol which was quantified by spectrophotometry at an absorbance of 420 nm. Two independent co-transformants (X and Y) are shown. As a negative control, the LexA-RAD51 expression vector was co-transformed with the pYesTrp2 vector (empty vector: EV). The values represent the mean of three independent experiments with the 95 % confidence intervals indicated by error bars.

This yeast two-hybrid analysis indicates that there is self-interaction between RAD51 molecules, which is consistent with published data from other eukaryotes, such as mammals (Benson *et al.*, 1994; Schild *et al.*, 2000)

### 5.4.2 Yeast two-hybrid analysis with LexA-RAD51-3 as the DNA-binding domain fusion

The next protein used as the “bait” was RAD51-3 and its interactions were assayed with RAD51 and the RAD51 paralogues as the “prey”. The vector pHybLex/Zeo-RAD51-3, expressing RAD51-3 as a LexA fusion, was co-transformed into *S. cerevisiae* L40 cells with pYesTrp2-RAD51, pYesTrp2-RAD51-3, pYesTrp2-RAD51-4, pYesTrp2-RAD51-5 and pYesTrp2-RAD51-6, expressing RAD51 and the RAD51 paralogues as V5-NLS-B42 fusions (see Materials and Methods, Section 2.4.6). As before, two independent co-transformants were analysed for each LexA-RAD51-3 fusion (X and Y) in combination with V5-NLS-B42-RAD51 and the V5-NLS-B42-RAD51 paralogue fusions. Figure 5-8A shows the western blot analysis of the co-transformants, demonstrating that each expressed

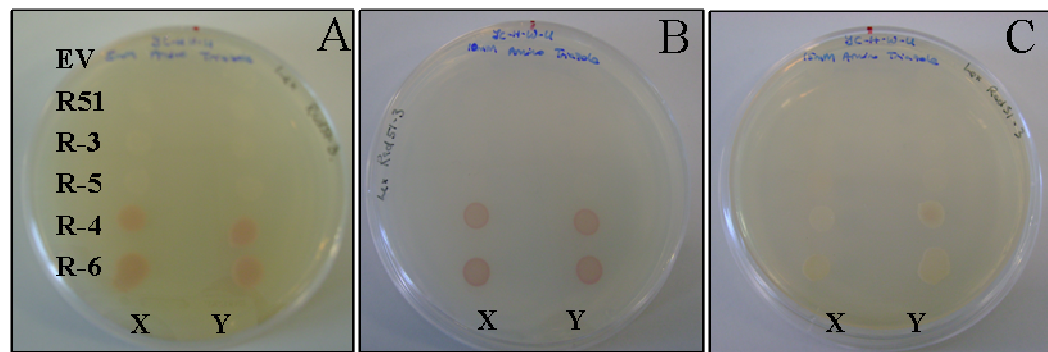
a protein of the predicted size of LexA-RAD51-3 (86 kDa) detected by the anti-LexA anti serum. This LexA-RAD51-3 fusion was absent in the untransformed *S. cerevisiae* L40 wildtype cell extract. Proteins of the expected sizes of the V5-NLS-B42 fusions to RAD51 and the RAD51 paralogues were also detected by the anti-V5 anti-serum (see Figure 5-8B).



**Figure 5-8: Western analysis of LexA-RAD51-3 expression with V5-NLS-B42-RAD51 and RAD51 paralogues in *S. cerevisiae* L40 co-transformants.**

Two co-transformants (X and Y) expressing LexA-RAD51-3 and each V5-NLS-B42 fusions of RAD51 (R51), RAD51-3 (R-3), RAD51-5 (R-5), RAD51-6 (R-6) or RAD51-4 (R-4) are shown. **Figure 5-8A:** Protein extracts from the co-transformants, and from the untransformed *S. cerevisiae* L40 cells (wt) were separated by SDS-PAGE, blotted and probed with anti-LexA anti-serum. **Figure 5-8B:** The blot was then stripped and reprobed with anti-V5 anti-serum. The inferred sizes of the proteins (kDa) are shown. The predicted sizes of the fusion polypeptides are as follows: LexA-RAD51-3: 86.5 kDa and V5-NLS-B42 fusions- RAD51: 52 kDa, RAD51-3: 66.5 kDa, RAD51-4: 57.7 kDa, RAD51-5: 53.6 kDa and RAD51-6: 62.4 kDa. The fusion of V5- NLS-B42 adds 12 kDa and the fusion of LexA adds 32 kDa to the proteins.

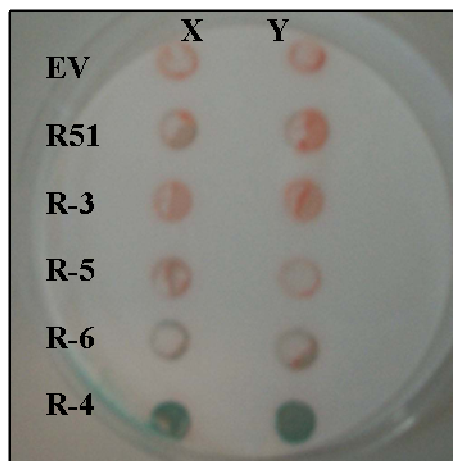
The results of the histidine auxotrophy test to examine interactions are shown in Figure 5-9, and reveal growth of the co-transformants expressing both LexA-RAD51-3 and each of V5-NLS-B42-RAD51-4 and V5-NLS-B42-RAD51-6 on media lacking histidine. This result indicates that RAD51-3 interacts with RAD51-4, and with RAD51-6. The absence of growth of the other co-transformants indicates that RAD51-3 does not interact with RAD51, RAD51-5 or with itself in this system.



**Figure 5-9: Histidine auxotrophy analysis of LexA-RAD51-3 with V5-NLS-B42-RAD51 and the V5-NLS-B42-RAD51 paralogues.**

**Figure 5-9** shows growth or absence of growth of the LexA-RAD51-3 fusion with V5-NLS-B42-RAD51 and each V5-NLS-B42-RAD51-paralogue fusion on medium lacking histidine. The two co-transformants (X and Y) are shown for LexA-RAD51-3 with each combination of V5-NLS-B42 fusion [Labelled: RAD51 (R51), RAD51-3 (R-3), RAD51-5 (R-5), RAD51-6 (R-6) or RAD51-4 (R-4)]. As a negative control, the LexA-RAD51-3 expression vector was co-transformed with the pYesTrp2 vector (empty vector: EV). The plates have three different concentrations of histidine inhibitor, 3' aminotriazole, **Figure 5-9A**: 5 mM, **B**: 10 mM, **C**: 15 mM.

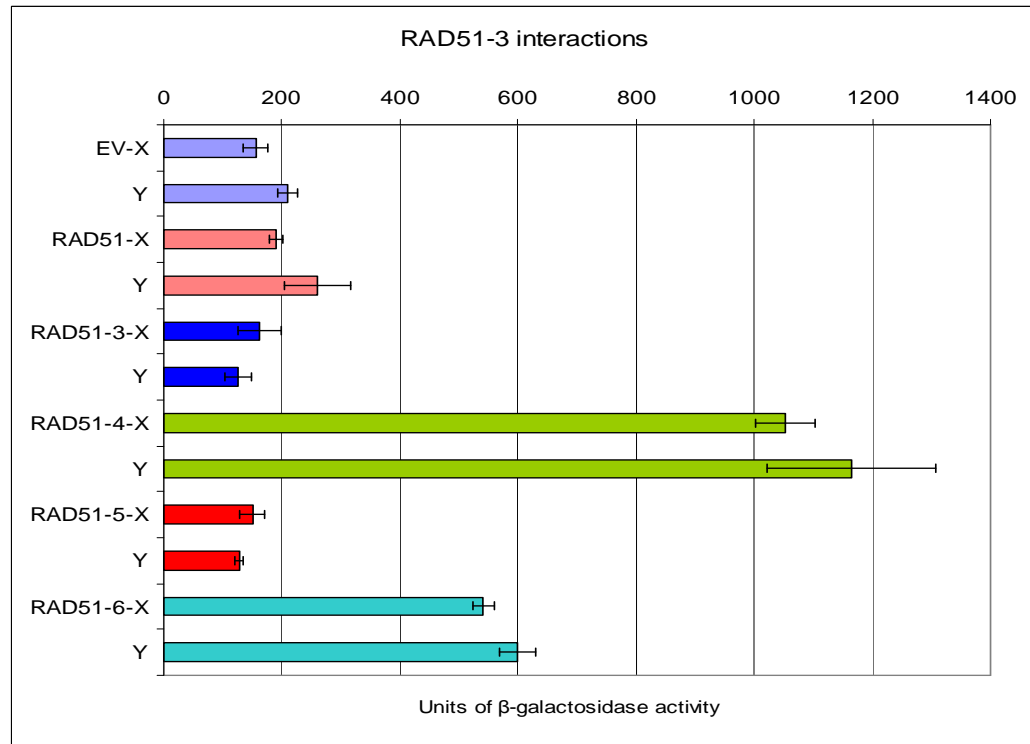
As for LexA-RAD51, the histidine auxotrophy assay of LexA-RAD51-3 interaction was verified by assaying for the  $\beta$ -galactosidase activity, firstly by the X-gal lift assay (see Figure 5-10), and secondly by quantifying the  $\beta$ -galactosidase activity in the yeast lysates (see Figure 5-11). A strong colour change from pink to blue was seen in the LexA-RAD51-3:V5-NLS-B42-RAD51-4 co-transformant in the X-gal filter lift assay (Figure 5-10), indicative of a relatively stable interaction. There was also evidence for a weak interaction between the LexA-RAD51-3 and V5-NLS-B42-RAD51-6 confirmed by colour change from pink to pale blue. No other interaction between LexA-RAD51-3 and V5-NLS-B42-RAD51, V5-NLS-B42-RAD51-3 or V5-NLS-B42-RAD51-5 could be detected with this assay.



**Figure 5-10: X-gal filter lift analysis of LexA-RAD51-3 with V5-NLS-B42-RAD51 and the V5-NLS-B42-RAD51 paralogues.**

As a positive control for the histidine auxotrophy test, the cultures of each co-transformant (X and Y) were plated on medium containing all the essential amino acids. These transformants were then analysed using the X-gal filter lift assay. LexA-RAD51-3 was co-expressed with each combination of V5-NLS-B42 fusion [Labelled: RAD51 (R51), RAD51-3 (R-3), RAD51-5 (R-5), RAD51-6 (R-6) or RAD51-4 (R-4)]. As a negative control, the LexA-RAD51-3 expression vector was co-transformed with the pYesTrp2 vector (empty vector: EV). Interactions are identified by a colour change from pale pink to blue, where  $\beta$ -galactosidase expression was induced from a *lacZ* reporter gene.

These weak and strong interactions were confirmed by quantitatively measuring  $\beta$ -galactosidase activity. The LexA-RAD51-3:V5-NLS-B42-RAD51-4 co-transformants displayed ~5-6 fold greater  $\beta$ -galactosidase activity than the LexA-RAD51-3 expressor control. The LexA-RAD51-3:V5-NLS-B42-RAD51-6 co-transformants also displayed an increase of  $\beta$ -galactosidase activity, but this was only ~3 fold higher than the control (see Figure 5-11). The data also confirm that RAD51-3 does not interact with RAD51, RAD51-3 or RAD51-5 in the yeast two-hybrid system.



**Figure 5-11: Quantification of β-galactosidase activity of LexA-RAD51-3 with V5-NLS-B42-RAD51 and the V5-NLS-B42-RAD51 paralogues.**

β-galactosidase activity was analysed for cells expressing LexA-RAD51-3 fusion with V5-NLS-B42-RAD51 and each V5-NLS-B42-RAD51-paralogue fusion using the yeast β-galactosidase assay kit by Pierce. The measurement of β-galactosidase activity considers the generation of *o*-nitrophenol which was quantified by spectrophotometry at an absorbance of 420 nm. Two independent co-transformants (X and Y) are shown. As a negative control, the LexA-RAD51-3 expression vector was co-transformed with the pYesTrp2 vector (empty vector: EV). The values represent the mean of three independent experiments with the 95 % confidence intervals indicated by error bars.

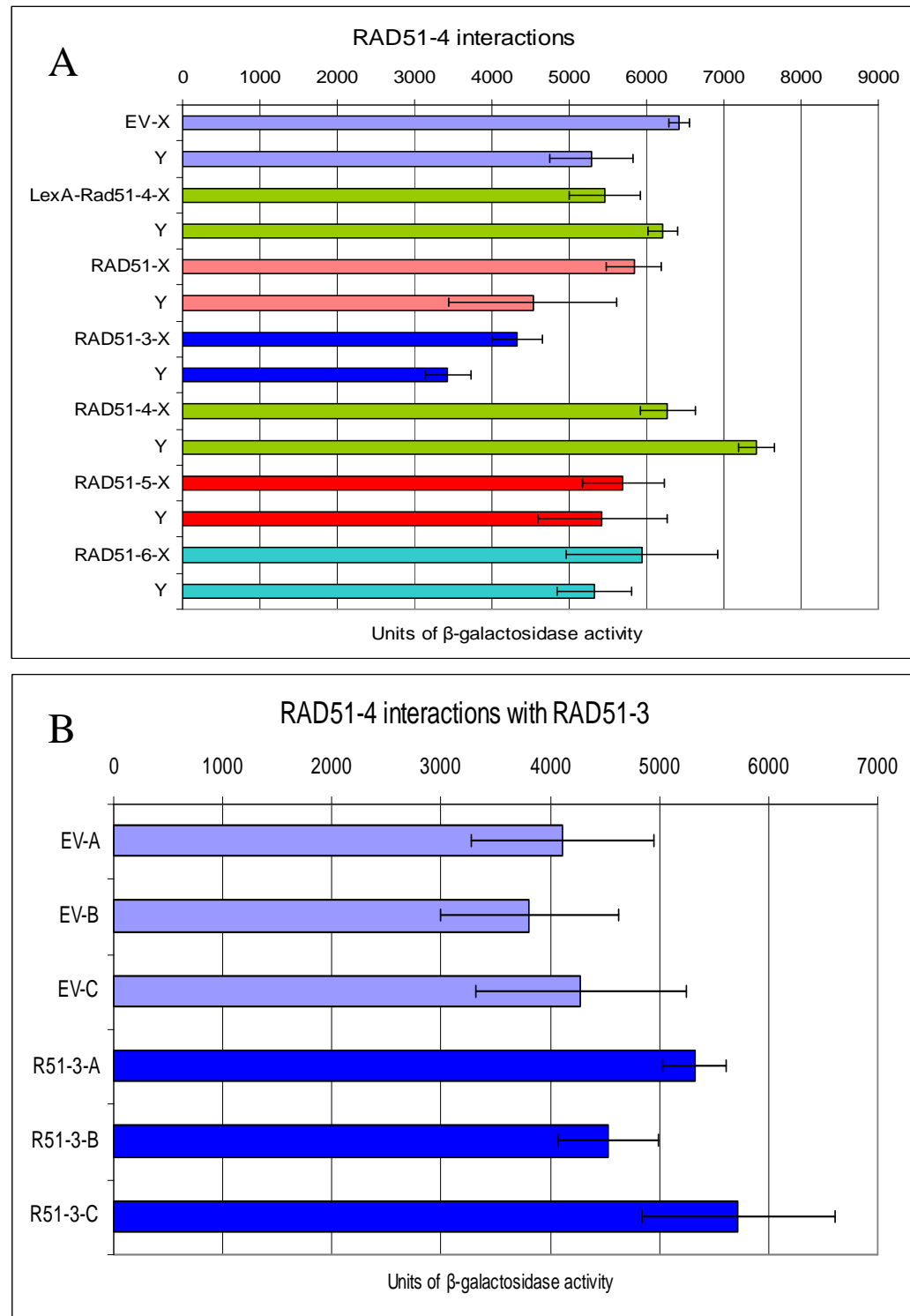
The results of all three assays are therefore consistent, showing an interaction between RAD51-3 and RAD51-4, and an interaction between RAD51-3 and RAD51-6.

### 5.4.3 Yeast two-hybrid analysis with LexA-RAD51-4 as the DNA-binding domain fusion

RAD51-4 was the next LexA fusion to be expressed to test for interactions with RAD51 and the other RAD51 paralogues. As before, all pair-wise combinations of pHybLex/Zeo-RAD51-4, expressing LexA-RAD51-4, and pYesTrp2-RAD51 and the four pYesTrp2-RAD51 paralogue vectors expressing V5-NLS-B42 fusions, were co-transformed into *S. cerevisiae* L40 cells (see Materials and Methods, Section 2.4.6). In this case, the LexA-RAD51-4 fusion appeared to act as a transcriptional activator, independently of the V5-NLS-B42 fusions. This is illustrated in Figure 5-12A, which show the β-galactosidase activity of LexA-RAD51-4 co-transformed with V5-NLS-B42-RAD51 and the V5-NLS-B42-RAD51 paralogue fusions. In all cases, including where LexA-RAD51-4 is expressed alone, very high levels of β-galactosidase activity were detected (~15-30 fold higher than

in the previous LexA-RAD51 and LexA-RAD51-3 experiments, which showed ~200 units of  $\beta$ -galactosidase activity; Figure 5-7 and Figure 5-11 respectively).

However, there was an apparent decrease of  $\beta$ -galactosidase activity when LexA-RAD51-4 was co-expressed with V5-NLS-B42-RAD51-3. It is conceivable that this indicates that LexA-RAD51-4 was bound by V5-NLS-B42-RAD51-3 in these conditions, sequestering the protein and resulting in less LexA-RAD51-4-mediated auto-induction of transcription by the B42 domain of the RAD51-3 fusion. To test this hypothesis, three further independent co-transformants expressing LexA-RAD51-4 and V5-NLS-B42-RAD51-3 were analysed (Figure 5-12B). In this case no apparent decrease in  $\beta$ -galactosidase activity was observed compared with LexA-RAD51-4 co-transformed with empty V5-NLS-B42 vector. This suggests that the potential interaction was not reproducible. Furthermore, the auto-induction catalysed by the full length RAD51-4 prevents interpretation of potential interactions in this yeast two-hybrid analysis.

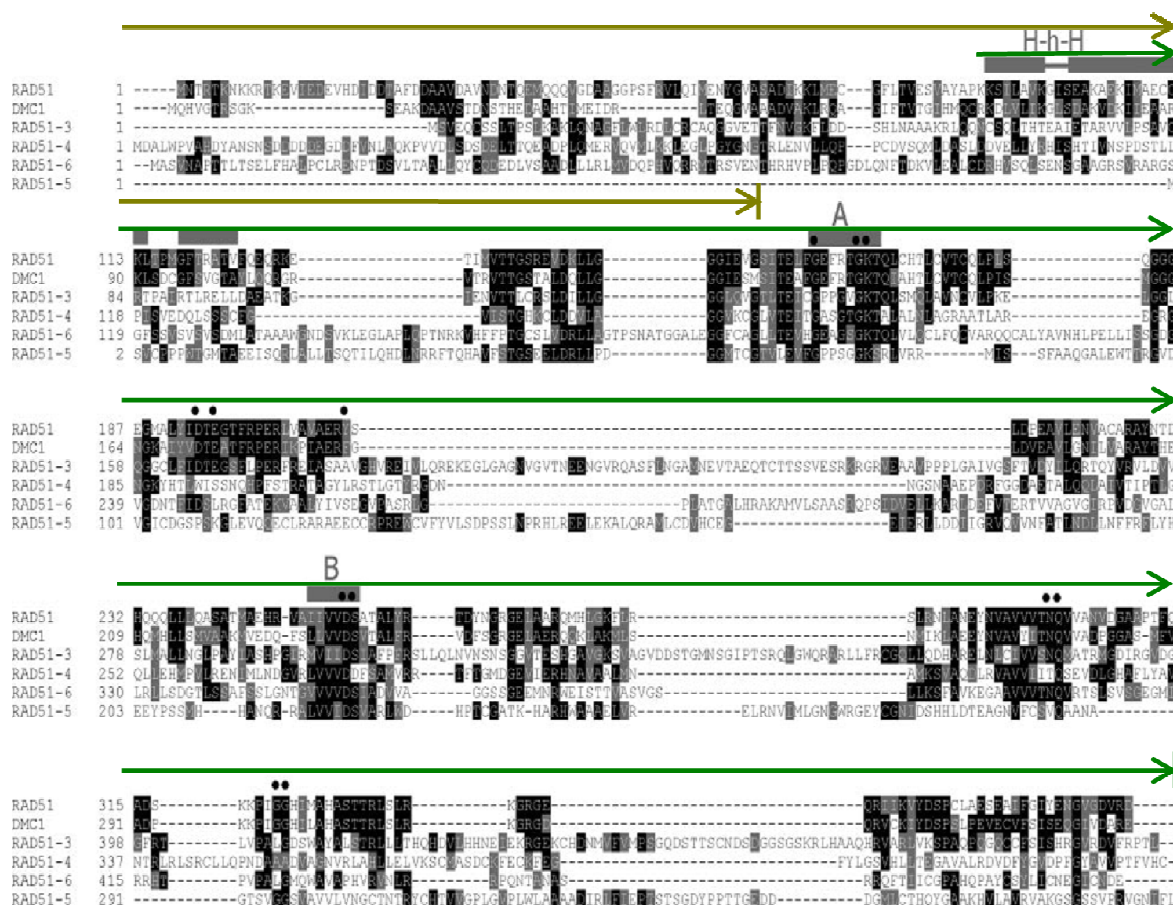


**Figure 5-12: Quantification of  $\beta$ -galactosidase activity of LexA-RAD51-4 co-expressed with pYesTrp2, V5-NLS-B42-RAD51 and each V5-NLS-B42-RAD51-paralogue fusion**

$\beta$ -galactosidase activity was analysed for cells expressing the LexA-RAD51-4 fusion with V5-NLS-B42-RAD51 and each V5-NLS-B42-RAD51-paralogue fusion assayed using the yeast  $\beta$ -galactosidase assay kit. **Figure 5-12A** shows the auto-induction of the LexA-RAD51-4 fusion with the V5-NLS-B42-RAD51 and the V5-NLS-B42-RAD51 paralogue fusions. As a negative control, the LexA-RAD51-4 expression vector was co-transformed with the pYesTrp2 vector [Labelled EV: empty vector]. As a further control, the LexA-RAD51-4 expression vector alone was analysed [Labelled LexA-RAD51-4]. The plot shows the values representing the mean of three independent experiments with the 95 % confidence intervals indicated by error bars. **Figure 5-12B** shows the auto-induction of LexA-RAD51-4 co-transformed with the pYesTrp2 vector (EV) and three independent LexA-RAD51-4:V5-NLS-B42-RAD51-3 transformants (A-C).

One possible reason for the auto-induction by LexA-RAD51-4 is that the sequence of the N terminus of RAD51-4 is rich in acidic amino acid residues; previous studies have shown that acidic sequences with no homology to proteins encoded elsewhere in the yeast genome can act as transcription activators (Ma & Ptashne, 1987b). To test this and to see if RAD51-4 interactions could be mapped, DNA encoding the C and N terminal fragment of the RAD51-4 protein were cloned into pHybLex/Zeo, based on the domains mapped in the human Rad51 paralogues (Miller *et al.*, 2004). Miller *et al.*, (2004) showed by yeast two-hybrid analysis, that the N terminal domain and the linker region of HsRad51B were required for interaction with the HsRad51C. The authors also showed that the N terminal domain of HsRad51B without the linker region could not interact with HsRad51C. Although highly speculative, and at present there is no direct evidence that the functions of the *T. brucei* and mammalian Rad51 paralogues are the same, there are some similarities between the species. For instance, both groups of paralogues have been shown to function in HR, and also have been shown to interact as complexes (Masson *et al.*, 2001; Liu *et al.*, 2002; Section 5.4 & Section 5.6 of this chapter). There are also potential Walker A and B boxes in both sub-sets of proteins, which allows some deduction of different domains from their relative positions (Miller *et al.*, 2004; Proudfoot & McCulloch, 2005a). The hypothesis behind making the truncations of the RAD51-4 protein was based on the findings of Miller *et al.*, 2004, with the aid of their predicted ribbon diagram and model of topology of the HsRad51 paralogues. The N terminal fragment of RAD51-4 encompassed the first 151 amino acids, which included a putative N terminal domain as well as the predicted helix-hairpin-helix linker region (Proudfoot & McCulloch, 2005a; Figure 5-13 ). It was this region of the HsRad51B that was necessary for binding to HsRad51C (Miller *et al.*, 2004). If the interacting domain is similar to the mammalian Rad51B paralogue and if the LexA-RAD51-4 N terminal fragment did not auto-induce, this truncation would allow interpretation of interactions with the other RAD51 paralogues. The C terminal fragment of RAD51-4 encompassed the amino acids 94 to 423, which again included the predicted helix-hairpin-helix linker region and the C terminal domain with the predicted Walker A and B boxes (Figure 5-13). It was this region of the HsRad51C that was necessary for binding to HsRad51B, HsRad51D and XRCC3 (Miller *et al.*, 2004). This could also act as a control, if LexA-RAD51-4 N terminal domain was shown to interact with one or more of the RAD51 paralogues, and the LexA-RAD51-4 C terminal domain did not interact and did not auto-induce, it could be deduced that the N terminus of the proteins is required for interaction, like the mammalian paralogue, HsRad51B.

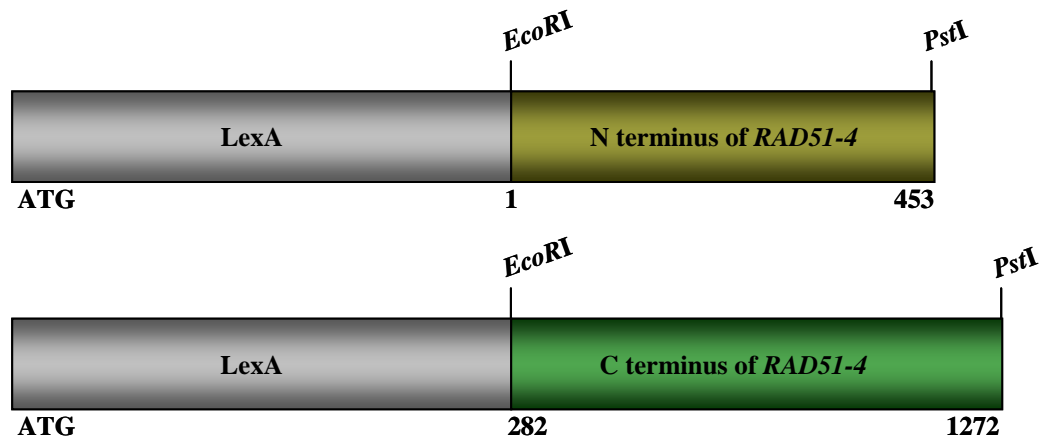




**Figure 5-13: A comparison of the RAD51 paralogue proteins highlighting the N and C terminal amino acid sequence domain used for truncation of RAD51-4.**

This alignment shows the predicted positions of the Walker A and B domains (A and B), and the helix-hairpin-helix (h-H-h) linker region. Also highlighted are the regions used for the RAD51-4 truncation: LexA-RAD51-4 N terminal domain is indicated with the brown arrow (1-151 amino acids) and the LexA-RAD51-4 C terminal domain is indicated with the green arrow (94-423 amino acids). Reprinted with permission from Oxford Journals, Proudfoot & McCulloch (2005a), *Nucleic Acid Research* **33**, pp 6909.

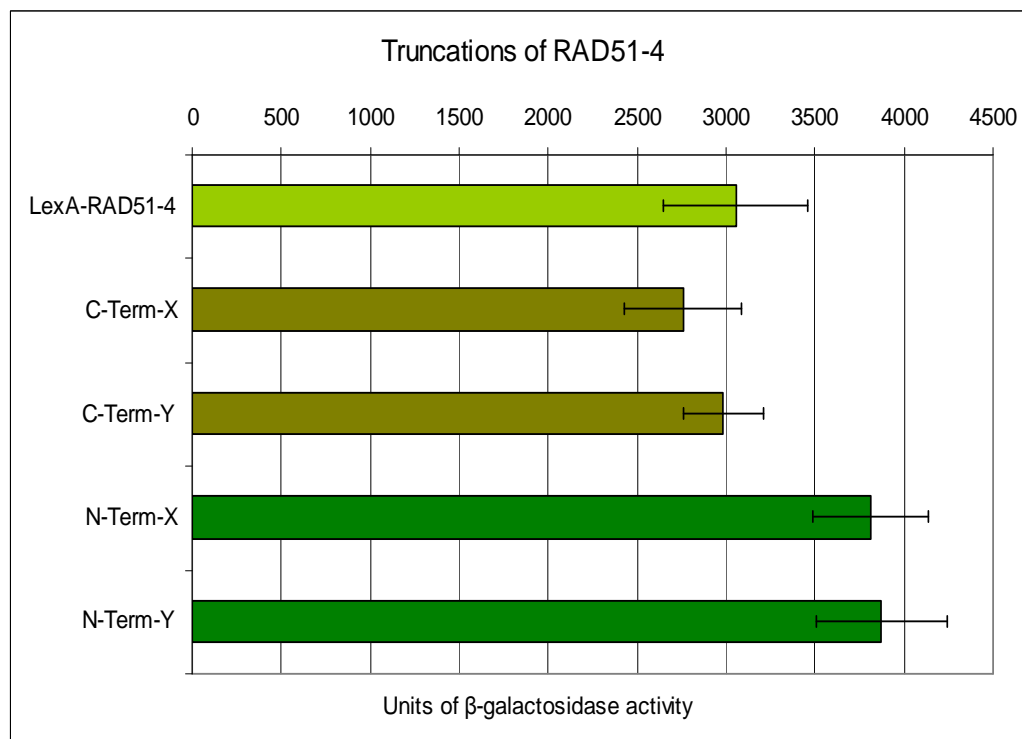
To this end, the coding sequence for the N terminus of RAD51-4 (from 1-453 bp) was PCR-amplified using primers N-TermR51-4 For and N-TermR51-4 Rev (see Materials and Methods, Section 2.4.1, Table 2-1). The resulting PCR fragment (453 bp) was then cloned into pHybLex/Zeo using the *EcoRI* and *PstI* sites, creating a LexA fusion to the first 1-151 amino acids of RAD51-4, as shown in Figure 5-14. The coding sequence for the C terminus of RAD51-4 (from 282-1272 bp) was PCR-amplified using primers C-TermR51-4 For and C-TermR51-4 Rev (see Materials and Methods, Section 2.4.1, Table 2-1). The 990 bp PCR-amplified fragment was then cloned in to pHybLex/Zeo using the *EcoRI* and *PstI* sites.



**Figure 5-14: Diagram showing the LexA fusions with the RAD51-4 truncations.**

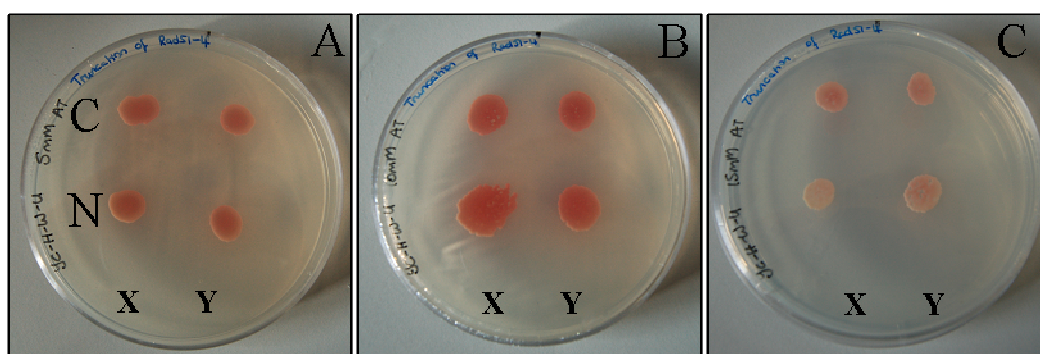
The complete *lexA* gene is part of the pHybLex/Zeo vector and allows expression of a fusion protein with the LexA DNA-binding domain. LexA has 202 amino acids, the N-terminal domain of RAD51-4 has 151 amino acids and the C-terminal domain of the RAD51-4 has 330 amino acids.

To determine if these fusions activated the expression of *lacZ* and *HIS3* genes, each construct, LexA-RAD51-4-N term and LexA-RAD51-4-C term, was co-transformed with the V5-NLS-B42 empty vector. Two co-transformants of each (X and Y) were examined. Both the  $\beta$ -galactosidase activity measurement (Figure 5-15) and histidine auxotrophy tests (Figure 5-16) showed that the truncations of RAD51-4 were also auto-inducers, resulting in equivalent levels of induction compared to the full length RAD51-4. The basis of this induction is unclear, but may reside in the 57 amino acids that overlap the two truncations. Nevertheless, this was not dissected further and evaluation of the LexA-RAD51-4 interactions by yeast two-hybrid cannot be interpreted.



**Figure 5-15: Auto-induction of the LexA-RAD51-4 C and N termini truncated fusions co-expressed with pYesTrp2 (EV) assayed using the yeast  $\beta$ -galactosidase assay kit.**

$\beta$ -galactosidase activity was analysed for cells expressing LexA-RAD51-4 C and N termini truncated fusions co-expressed with the pYesTrp2 vector expressing only V5-NLS-B42. Auto-induction was apparent for both the independent co-transformants (X and Y) for LexA-RAD51-4 C and LexA-RAD51-4 N termini truncated fusions. As a control, the LexA-RAD51-4 expression vector alone showed auto-induction [Labelled LexA-RAD51-4]. The plot shows the values representing the mean of three independent experiments with the 95 % confidence intervals indicated by error bars.



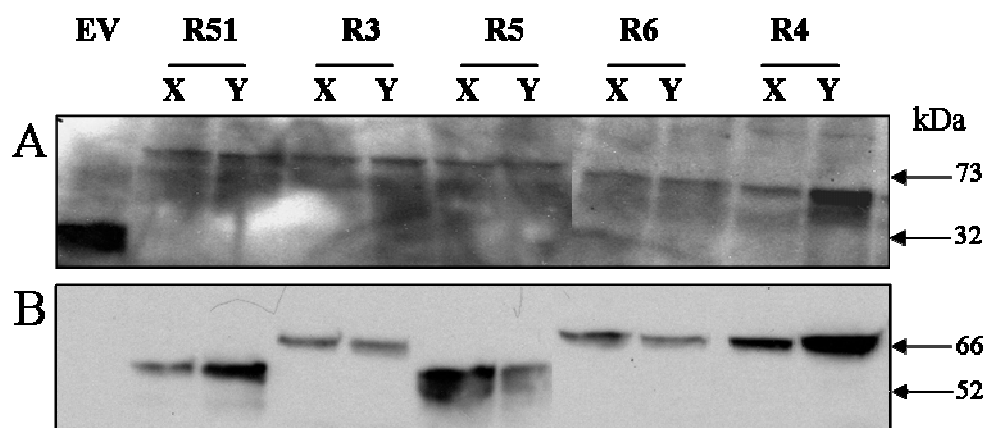
**Figure 5-16: Auto-induction of the truncated LexA-RAD51-4 C and N termini truncated fusions co-expressed with pYesTrp2 (EV) assayed by histidine auxotrophy.**

**Figure 5-16** shows the growth of the LexA-RAD51-4 C and N terminal fusion without a V5-NLS-B42 fusion protein, on medium lacking histidine. Two independent co-transformants (X and Y) are shown for each of the LexA-RAD51-4 C and N termini truncated fusions. The plates have three different concentrations of histidine inhibitor, 3' aminotriazole, **Figure 5-16A**: 5 mM, **B**: 10 mM, **C**: 15 mM.

### 5.4.4 Yeast two-hybrid analysis with LexA-RAD51-5 as the DNA-binding domain fusion

The fourth protein expressed as a LexA fusion to examine interaction was RAD51-5. The pHybLex/Zeo-RAD51-5 construct was generated by PCR-amplification of the RAD51-5 ORF using the primers R51-5 For and R51-5 Rev (see Materials and Methods, Section 2.4.1, Table 2-1). This fragment was then cloned into pHybLex/Zeo using flanking *ApaI* and *SalI* sites (Figure 5-13). This strategy resulted in a frame shift, and therefore meant that the LexA and RAD51-5 ORFs were out of frame. Site directed mutagenesis was carried out to rectify this: an extra nucleotide (A) was added to the *ApaI* site, resulting in disruption of the *ApaI* restriction site, which allowed possible mutated clones to be screened. This also resulted in correction of the frame shift, making the RAD51-5 ORF in frame with LexA; the corrected construct was verified by DNA sequencing.

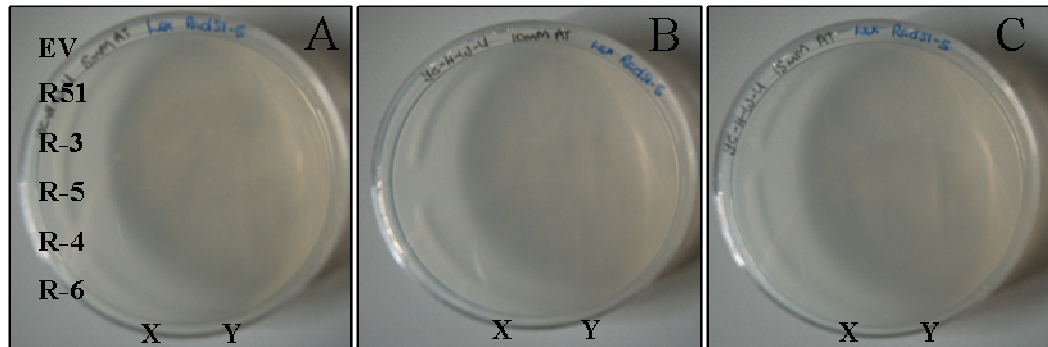
pHybLex/Zeo-RAD51-5 expressing a LexA-RAD51-5 fusion, was co-expressed with pYesTrp2-RAD51 and pYesTrp2-RAD51 paralogues as V5-NLS-B42 fusions in *S. cerevisiae* L40 cells (see Materials and Methods, Section 2.4.6). Two independent co-transformants were analysed for each combination (X and Y). LexA-RAD51-5 fusion expression was verified by western blotting using anti-LexA anti-serum (Figure 5-17A) and the appropriately sized V5-NLS-B42 fusions were found to be expressed in the cognate co-transformants by detection with anti-V5 anti-serum (Figure 5-17B).



**Figure 5-17: Western analysis of LexA-RAD51-5 expression with V5-NLS-B42-RAD51 and RAD51 paralogues in *S. cerevisiae* L40 co-transformants.**

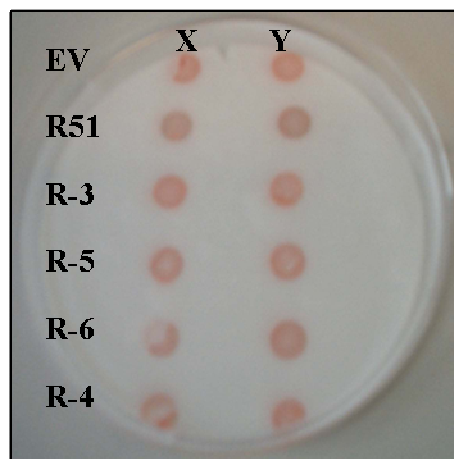
Two co-transformants (X and Y) expressing LexA-RAD51-5 and each V5-NLS-B42 fusion of RAD51 (R51), RAD51-3 (R-3), RAD51-5 (R-5), RAD51-6 (R-6) or RAD51-4 (R-4) are shown. **Figure 5-17A:** Protein extracts from the co-transformants were separated by SDS-PAGE, blotted and probed with anti-LexA anti-serum. The predicted size of LexA-RAD51-5 is 73.6 kDa. The negative control (EV) shows a pHybLex/Zeo empty vector transformant and indicates the predicted size of the non-fused LexA: 32 kDa. **Figure 5-17B:** The blot was then stripped and reprobed with anti-V5 anti-serum. The inferred sizes of the proteins (kDa) are shown. The predicted sizes of the fusion polypeptides are as follows: V5-NLS-B42 fusions- RAD51: 52 kDa, RAD51-3: 66.5 kDa, RAD51-4: 57.7 kDa, RAD51-5: 53.6 kDa and RAD51-6: 62.4 kDa. The fusion of V5-NLS-B42 adds 12 kDa.

The histidine auxotrophy test showed no evidence for transcriptional activation when the LexA-RAD51-5 fusion was co-expressed with V5-NLS-B42-RAD51 or any of the V5-NLS-B42-RAD51 paralogues (Figure 5-18). The lack of observable interaction was confirmed by assaying for  $\beta$ -galactosidase activity, using both the X-gal filter lift assay (Figure 5-19) and measuring the  $\beta$ -galactosidase activity in cell lysates (Figure 5-20).



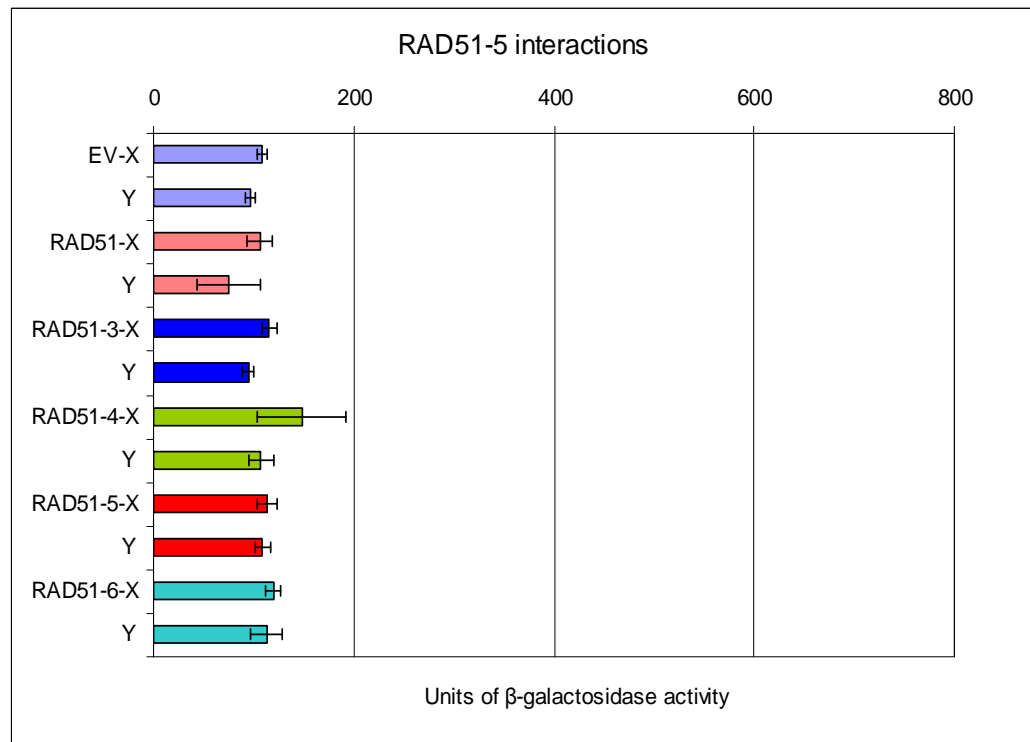
**Figure 5-18: Histidine auxotrophy analysis of LexA-RAD51-5 with V5-NLS-B42-RAD51 and the V5-NLS-B42-RAD51 paralogues.**

The diagram shows growth or absence of growth of the LexA-RAD51-5 fusion with V5-NLS-B42-RAD51 and each V5-NLS-B42-RAD51-paralogue fusion on medium lacking histidine. The two co-transformants (X and Y) are shown for LexA-RAD51-5 with each combination of V5-NLS-B42 fusion [Labelled: RAD51 (R51), RAD51-3 (R-3), RAD51-5 (R-5), RAD51-6 (R-6) or RAD51-4 (R-4)]. As a negative control, the LexA-RAD51-5 expression vector was co-transformed with the pYesTrp2 vector (empty vector: EV). The plates have three different concentrations of histidine inhibitor, 3' aminotriazole, **Figure 5-18A**: 5 mM, **B**: 10 mM, **C**: 15 mM.



**Figure 5-19: X-gal filter lift analysis of LexA-RAD51-5 with V5-NLS-B42-RAD51 and the V5-NLS-B42-RAD51 paralogues.**

As a positive control for the histidine auxotrophy test, the cultures of each co-transformant (X and Y) were plated on medium containing all the essential amino acids. These transformants were then analysed using the X-gal filter lift assay. LexA-RAD51-5 was co-expressed with each combination of V5-NLS-B42 fusion [Labelled: RAD51 (R51), RAD51-3 (R-3), RAD51-5 (R-5), RAD51-6 (R-6) or RAD51-4 (R-4)]. As a negative control, the LexA-RAD51-5 expression vector was co-transformed with the pYesTrp2 vector (empty vector: EV). Interactions are identified by a colour change from pale pink to blue, where  $\beta$ -galactosidase expression was induced from a *lacZ* reporter gene.



**Figure 5-20: Quantification of β-galactosidase activity of LexA-RAD51-5 with V5-NLS-B42-RAD51 and the V5-NLS-B42-RAD51 paralogues.**

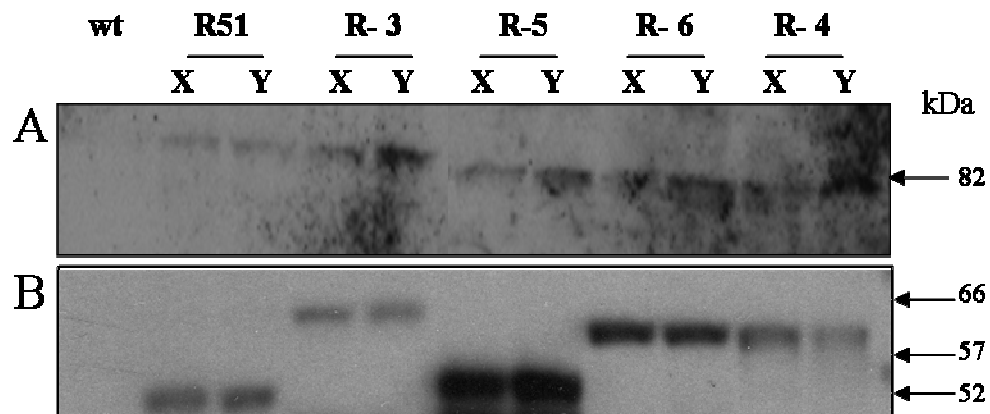
β-galactosidase activity was analysed for cells expressing the LexA-RAD51-5 fusion with V5-NLS-B42-RAD51 and each V5-NLS-B42-RAD51-paralogue fusion using the yeast β-galactosidase assay kit by Pierce. The measurement of β-galactosidase activity considers the generation of *o*-nitrophenol which was quantified by spectrophotometry at an absorbance of 420 nm. Two independent co-transformants (X and Y) are shown. As a negative control, the LexA-RAD51-5 expression vector was co-transformed with the pYesTrp2 vector (empty vector: EV). The values represent the mean of three independent experiments with the 95 % confidence intervals indicated by error bars.

Since the western analysis shows that the LexA-RAD51-5 fusion is detectably expressed, as are the V5-NLS-B42-RAD51 and the V5-NLS-B42-RAD51 paralogues fusions (Figure 5-17), the lack of reporter gene expression is not simply explained by poor expression or rapid degradation of the proteins. These results may indicate that RAD51-5 may not interact with any other proteins as a “bait” in this assay. It may of course be the case that the LexA-RAD51-5 fusion is not localised to the nucleus, or the fusion cannot bind to the DNA. It could also be the case that the fusion to LexA causes incorrect folding of the RAD51-5 protein, disrupting the interaction binding sites. Nevertheless, these results may indicate that RAD51-5 does not interact with the other *T. brucei* RAD51 paralogues.

#### 5.4.5 Yeast two-hybrid analysis with LexA-RAD51-6 as the DNA-binding domain fusion

Finally, RAD51-6 was used as the “bait” for interaction analysis. The pHybLex/Zeo-RAD51-6 was co-transformed with all pair-wise combinations of pYesTrp2-RAD51 and each of the pYesTrp2-RAD51 paralogues into *S. cerevisiae* L40 cells (see Materials and

Methods, Section 2.4.6). As before, two independent co-transformants (X and Y) were analysed for expression of the LexA-RAD51-6 fusion and for the cognate V5-NLS-B42-RAD51 or V5-NLS-B42-RAD51 paralogue fusions (Figure 5-21). LexA-RAD51-6 expression in each co-transformant was identified: although weak, a band of the predicted size (~83 kDa) was detected using anti-LexA anti-serum. The V5-NLS-B42 fusions to RAD51 and to the RAD51 paralogs were also readily detected, using anti-V5 anti-serum.

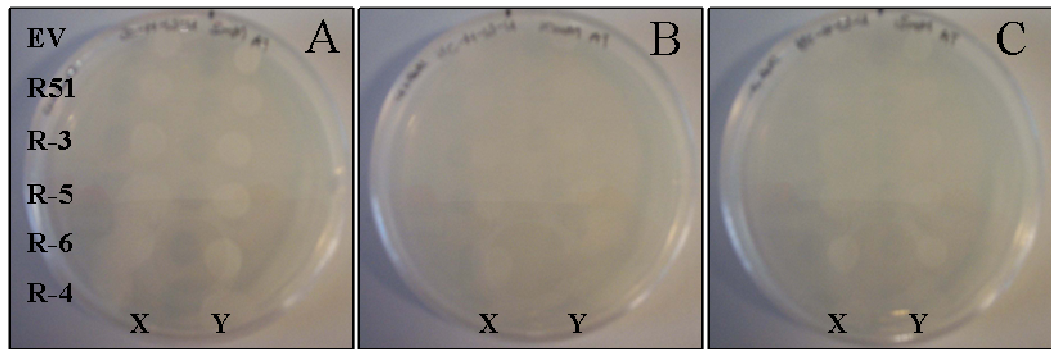


**Figure 5-21: Western analysis of LexA-RAD51-6 expression with V5-NLS-B42-RAD51 and RAD51 paralogs in *S. cerevisiae* L40 co-transformants.**

Two co-transformants (X and Y) expressing LexA-RAD51-6 and each V5-NLS-B42 fusion of RAD51 (R51), RAD51-3 (R-3), RAD51-5 (R-5), RAD51-6 (R-6) or RAD51-4 (R-4) are shown. **Figure 5-21A:** Protein extracts from the co-transformants, and for the untransformed *S. cerevisiae* L40 cells (wt) were separated by SDS-PAGE, blotted and probed with anti-LexA anti-serum. **Figure 5-21B:** The blot was then stripped and reprobed with anti-V5 anti-serum. The inferred sizes of the proteins (kDa) are shown. The predicted sizes of the fusion polypeptides are as follows: LexA-RAD51-6: 82.4 kDa and V5-NLS-B42 fusions- RAD51: 52 kDa, RAD51-3: 66.5 kDa, RAD51-4: 57.7 kDa, RAD51-5: 53.6 kDa and RAD51-6: 62.4 kDa. The fusion of V5-NLS-B42 adds 12 kDa and the fusion of LexA adds 32 kDa to the proteins.

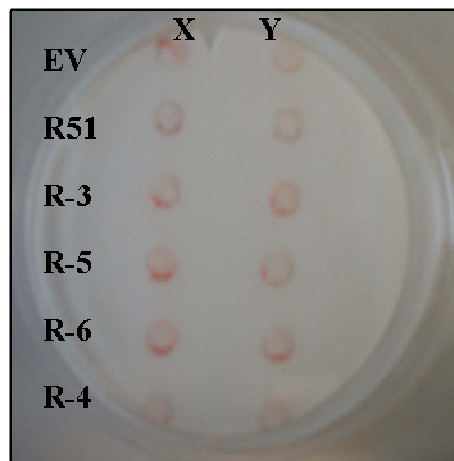
The histidine auxotrophy test revealed no evidence for transcriptional activation when the LexA-RAD51-6 fusion was co-expressed with V5-NLS-B42-RAD51 or the other V5-NLS-B42-RAD51 paralogs (Figure 5-22). This negative result was confirmed using with the X-gal filter lift assay (Figure 5-23) to detect  $\beta$ -galactosidase activity and was also confirmed by quantification of  $\beta$ -galactosidase activity in the lysates of the co-transformants (Figure 5-24).





**Figure 5-22: Histidine auxotrophy analysis of LexA-RAD51-6 with V5-NLS-B42-RAD51 and the V5-NLS-B42-RAD51 paralogues.**

**Figure 5-22** shows growth or absence of growth of the LexA-RAD51-6 fusion with V5-NLS-B42-RAD51 and each V5-NLS-B42-RAD51-paralogue fusion on medium lacking histidine. The two co-transformants (X and Y) are shown for LexA-RAD51-6 with each combination of V5-NLS-B42 fusion [Labelled: RAD51 (R51), RAD51-3 (R-3), RAD51-5 (R-5), RAD51-6 (R-6) or RAD51-4 (R-4)]. As a negative control, the LexA-RAD51-6 expression vector was co-transformed with the pYesTrp2 vector (empty vector: EV). The plates have three different concentrations of histidine inhibitor, 3' aminotriazole, **Figure 5-22A:** 5 mM, **B:** 10 mM, **C:** 15 mM.



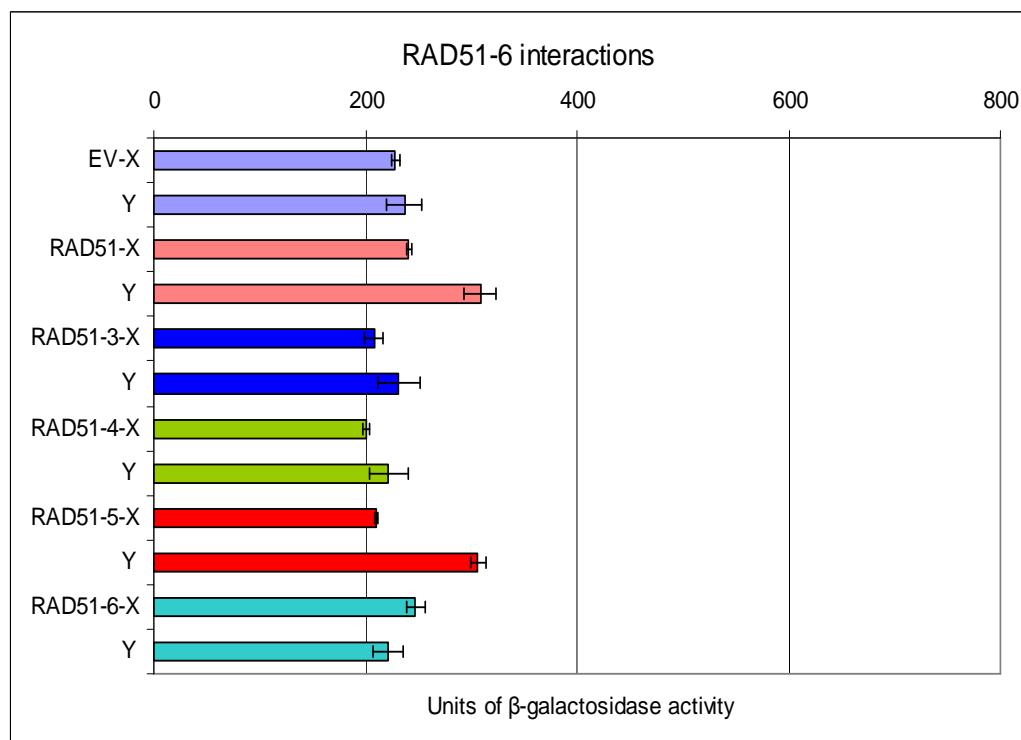
**Figure 5-23: X-gal filter lift analysis of LexA-RAD51-6 with V5-NLS-B42-RAD51 and the V5-NLS-B42-RAD51 paralogues.**

As a positive control for the histidine auxotrophy test, the cultures of each co-transformant (X and Y) were plated on medium containing all the essential amino acids. These transformants were then analysed using the X-gal filter lift assay. LexA-RAD51-6 was co-expressed with each combination of V5-NLS-B42 fusion [Labelled: RAD51 (R51), RAD51-3 (R-3), RAD51-5 (R-5), RAD51-6 (R-6) or RAD51-4 (R-4)]. As a negative control, the LexA-RAD51-6 expression vector was co-transformed with the pYesTrp2 vector (empty vector: EV). Interactions are identified by a colour change from pale pink to blue, where  $\beta$ -galactosidase expression was induced from a *lacZ* reporter gene.

Western analysis showed the LexA-RAD51-6 fusion is detectably expressed, as are the V5-NLS-B42 fusions in the co-transformants (Figure 5-21). It was previously shown that LexA-RAD51-3 interacted with V5-NLS-B42-RAD51-6 (Section 5.4.2). It therefore is surprising that the reciprocal experiment, using LexA-RAD51-6 fusion as the “bait” and the V5-NLS-B42-RAD51-3 fusion as the “prey” did not reveal any interaction. One possible explanation is that the C terminal fusion of LexA to RAD51-6 inhibits the site of interaction between RAD51-6 and RAD51-3. Another explanation is that one or both of



the fusion proteins may not be localised to the nucleus. V5-NLS-B42-RAD51-3 fusion is the largest of the fused proteins at 86.5 kDa and did not activate transcription in combination with any of the other LexA fusion proteins. It is also possible that the B42 domain might be unable to recruit RNA polymerase when fused to RAD51-3 or when RAD51-6 is bound to the RAD51-3.



**Figure 5-24: Quantification of β-galactosidase activity of LexA-RAD51-6 with V5-NLS-B42-RAD51 and the V5-NLS-B42-RAD51 paralogues.**

β-galactosidase activity was analysed for cell expressing the LexA-RAD51-6 fusion with V5-NLS-B42-RAD51 and each V5-NLS-B42-RAD51-paralogue fusion using the yeast β-galactosidase assay kit by Pierce. The measurement of β-galactosidase activity considers the generation of *o*-nitrophenol which was quantified by spectrophotometry at an absorbance of 420 nm. Two independent co-transformants (X and Y) are shown. As a negative control, the LexA-RAD51-6 expression vector was co-transformed with the pYesTrp2 vector (empty vector: EV). The values represent the mean of three independent experiments with the 95 % confidence intervals indicated by error bars.

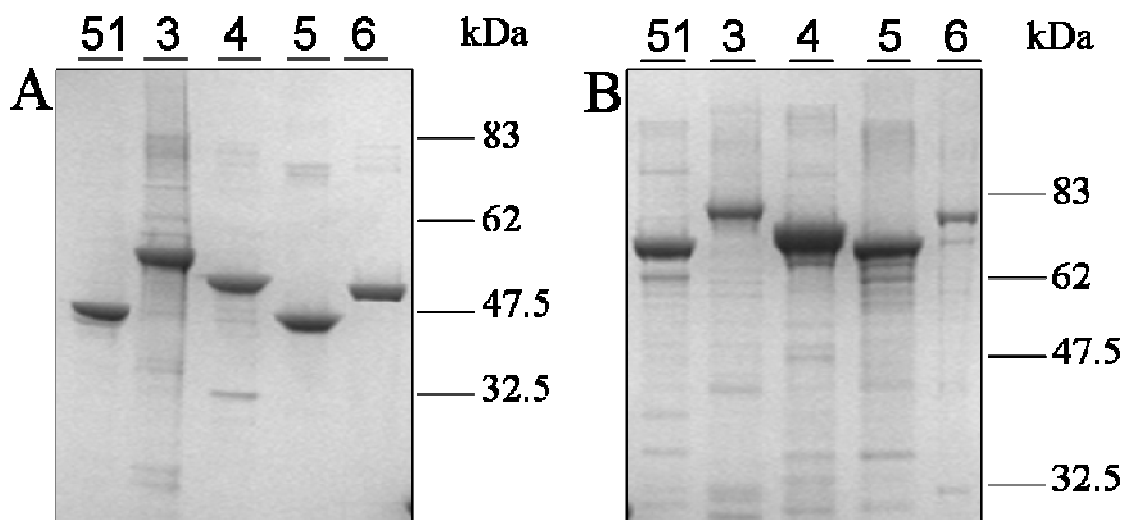
## 5.5 Introduction to the raising of anti-RAD51 and anti-RAD51 paralogue anti-sera

The section above attempts to determine if the *T. brucei* RAD51 paralogues interact with each other using yeast two-hybrid analysis. This showed that RAD51 interacts with itself and RAD51-3 interacts with RAD51-4 and RAD51-6, whilst no other interactions between RAD51 and the RAD51 paralogues were identified. In order to confirm that these interactions occur *in vivo*, and possibly to identify other interactions, co-immunoprecipitation was carried out. To this end, anti-sera were raised against RAD51 and the RAD51 paralogues which were then used in co-immunoprecipitation analysis from

*T. brucei* cell extracts. To describe this, Sections 5.5.1 to 5.5.5 below discuss the way these anti-sera were generated, tested and purified.

The anti-sera were raised against C-terminal 10 X HIS-tagged recombinant variants of RAD51 and each of the RAD51 paralogues. This purification was carried out by Dr Chris Stockdale, and involved cloning each ORF of *RAD51*, *RAD51-3*, *RAD51-4*, *RAD51-5*, and *RAD51-6* into the pET51b vector (Novagen), which allowed expression of the fusion proteins in *E. coli* BL21-rosetta2 (Novagen). Figure 5-25A shows a Coomassie stained SDS-PAGE gel of the different C-terminal HIS-tagged proteins after their purification by nickel column chromatography.

Each of these purified proteins were quantified and sent to Scottish National Blood Transfusion Service, Edinburgh for injection into different host animals for collection of total anti-serum. Rabbit was used to raise anti-serum against the RAD51 and RAD51-4 recombinant proteins, sheep was used to raise anti-serum against RAD51-3, chicken was used to raise anti-serum against RAD51-5, and rat was used to raise anti-serum against RAD51-6.



**Figure 5-25: C-terminal HIS-tagged and N-terminal GST-tagged recombinant RAD51 and the four RAD51 paralogues proteins**

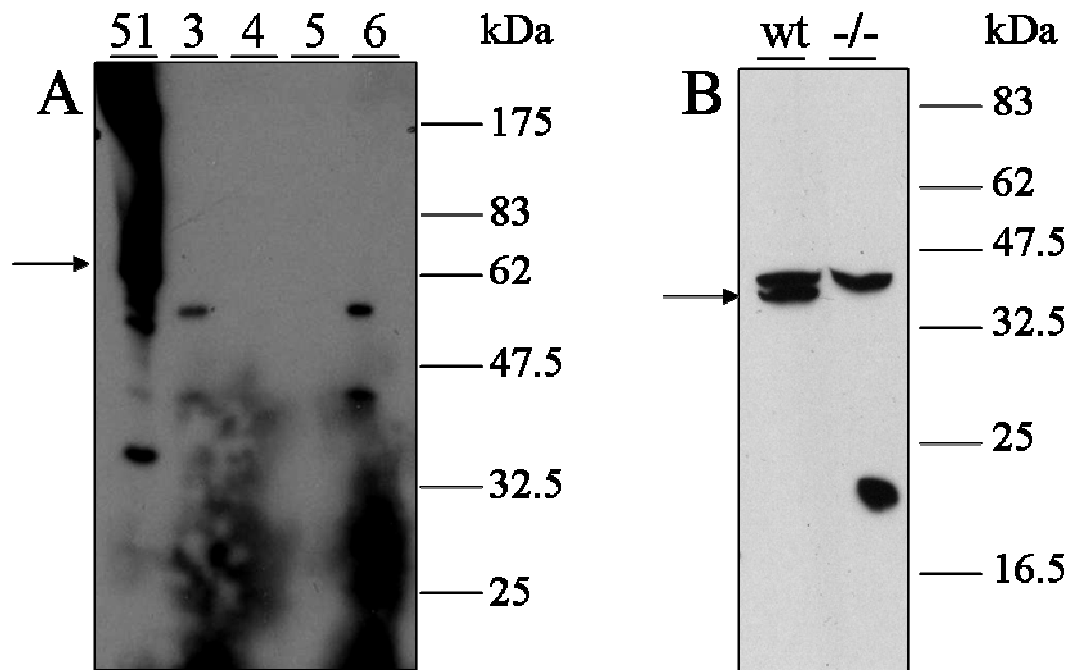
**Figure 5-25A:** Coomassie stained SDS-PAGE gel of an example of the separated purified proteins used to raise the anti-sera. Each protein was tagged with the C-terminal HIS: Labelled RAD51 [51], RAD51-3 [3], RAD51-4 [4], RAD51-5 [5] and RAD51-6 [6]. This 10 X HIS tag fusion added 2.5 kDa to RAD51 and the RAD51 paralogues. The predicted sizes of the fusion proteins are RAD51-HIS: 42.5 kDa; RAD51-3-HIS: 57 kDa; StrepII-RAD51-4-HIS: 54.6 kDa; RAD51-5-HIS: 44.1 kDa; and RAD51-6-HIS: 52.9 kDa. **Figure 5-25B:** Coomassie stained SDS-PAGE gel of an example of the separated purified proteins used to purify the anti-RAD51 and anti-RAD51 paralogue anti-sera. This GST tag fusion added 26.3 kDa to RAD51 and the RAD51 paralogues. The predicted sizes of the fusion proteins are GST-RAD51: 66.3 kDa; GST-RAD51-3: 80.3 kDa; GST-RAD51-4: 73.8 kDa; GST-RAD51-5: 67.9 kDa; and GST-RAD51-6: 76.7 kDa. These Figures were provided by Dr Chris Stockdale.

For anti-sera purification, N-terminal GST-tagged proteins were purified. Each ORF of the RAD51 paralogues were cloned into pGEX-4T3 (Amersham) and proteins purified from *E. coli* BL21-rosetta2 (Novagen) by Glutathione Affinity Chromatography; this work was completed by Dr Chris Stockdale. Figure 5-25B shows the Coomassie stained SDS-PAGE gel of an example of the separated GST-tagged purified proteins. These GST-tagged proteins were coupled to beads using Amino-Link (Pierce) and were then used as an affinity column for purification of the anti-sera. The affinity columns were made by Dr Chris Stockdale. Once made, the affinity column was stored at 4 °C until required.

### 5.5.1 Anti-RAD51 anti-serum purification

Specificity of the rabbit RAD51 anti-serum was analysed first. To do this, the RAD51 anti-serum was used to probe all of the GST-tagged RAD51 and RAD51 paralogue recombinant proteins (see Figure 5-25B). For this, a concentration of 2 ng.µl<sup>-1</sup> of each purified GST-tagged recombinant protein was separated on a 10 % SDS-PAGE gel. The proteins were transferred to a nylon membrane and probed with the rabbit anti-RAD51 anti-serum (at a dilution of 1:500) and then detected with anti-rabbit IgG-HRP (Molecular Probes) anti-serum (at a 1:5000 dilution). Figure 5-26A shows the western blot that was generated. The GST tag adds 26.3 kDa to the size of RAD51 and the RAD51 paralogue proteins. The black arrow indicates the predicted size of the GST-RAD51 protein (66.4 kDa). A very pronounced smear was seen in the lane containing GST-RAD51, possibly because of degradation of the recombinant protein. More likely, the anti-RAD51 serum bound to cross-contaminating proteins from *E. coli* during the GST-tag purification, as absolute purification is doubtful. These contaminants would also be targeted by the rabbit antibodies. Nevertheless, no bands of the predicted sizes in the other lanes for the GST-tagged RAD51-3, RAD51-4, RAD51-5 or RAD51-6 proteins were seen, indicating that the anti-RAD51 anti-sera did not bind to the GST-tagged RAD51 paralogues. Next, the specificity of anti-RAD51 anti-serum was tested against whole cell protein extracts from *T. brucei* cells. This was done by separating the extracted proteins from wildtype bloodstream stage Lister strain 427 and from bloodstream stage 3174.2 *rad51* homozygous mutant (-/-) on a 10 % SDS-PAGE gel; see Materials and Methods, Section 2.6.2.1 for protein extraction details and Section 2.6.3 for western blotting. The generation of the *rad51*<sup>-/-</sup> was described in McCulloch & Barry (1999) and Chapter 4 (Section 4.1). The separated extracts were then transferred to a nylon membrane, which was probed with the rabbit anti-RAD51 anti-serum (at a dilution of 1:500) and then detected with anti-rabbit IgG-HRP (Molecular Probes) anti-serum (at a 1:5000 dilution; Figure 5-26B). Figure 5-26B shows a band of the predicted size of RAD51 in the wildtype cell extract of 40.5 kDa, whilst this

band is absent in the *rad51*  $-/-$  protein extract. The blot shows a non-specific, cross-reactive band slightly larger than the predicted RAD51 band.



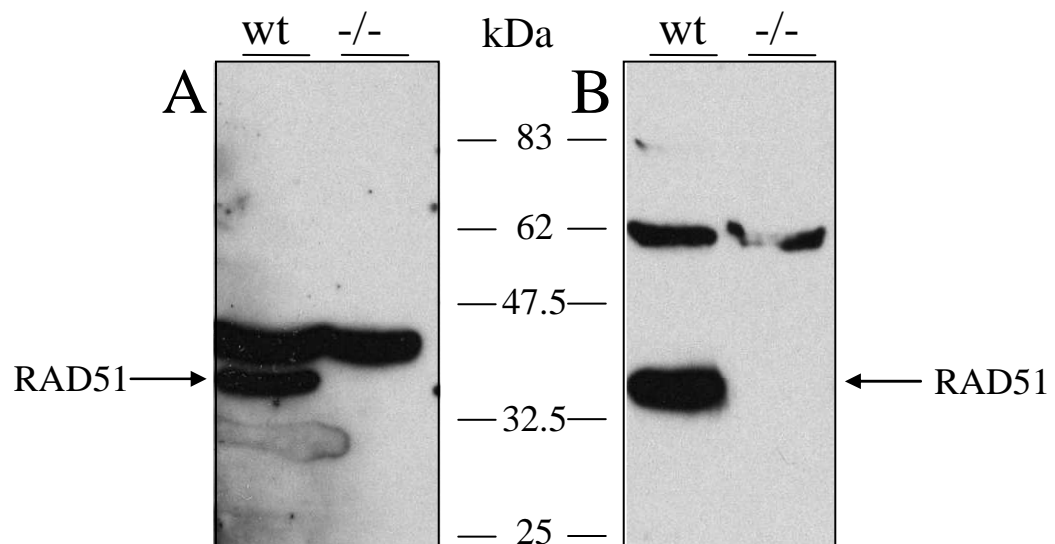
**Figure 5-26: Western analysis showing the specificity of the anti-RAD51 anti-serum.**

**Figure 5-26A:** Western blot shows the separated GST-RAD51 and GST-RAD51 paralogue recombinant proteins probed with anti-RAD51 anti-serum. [Labelled: RAD51 (51), RAD51-3 (3), RAD51-4 (4), RAD51-5 (5), and RAD51-6 (6)]. The GST tag adds 26.3 kDa to RAD51 and the RAD51 paralogs. The predicted sizes of the fusion proteins are GST-RAD51-3: 80.8 kDa, GST-RAD51-4: 72 kDa, GST-RAD51-5: 67.9 kDa and GST-RAD51-6: 76.7 kDa. The black arrow indicates the predicted size of the GST-RAD51 protein of 66.4 kDa. **Figure 5-26B:** Western blot shows protein extract from wildtype (wt) and *rad51* homozygous mutant ( $-/-$ ) *T. brucei* cells probed with the anti-RAD51 anti-sera. The black arrow indicates the predicted size of native RAD51 (40.5 kDa). The size markers are indicated for both blots.

As this cross-reactive band may have interfered with further analysis, antibody purification from this anti-serum was performed. To do this, the purified GST-RAD51 protein (Figure 5-25) was coupled to beads using Amino-link (Pierce) which were then used as an affinity column. Approximately 4 ml of anti-serum were added to the GST-RAD51 affinity column, and incubated for 1 hour to allow the anti-RAD51 antibodies to bind to the GST-RAD51 (see Materials and Methods, Section 2.7.2). This was then washed thoroughly with PBS, with unbound or weakly bound proteins discarded in the flow-through. The purified anti-RAD51 antibodies were then eluted from the GST-RAD51 bound beads using 100 mM glycine-HCl (pH 2.7) solution, in fractions of 0.5 ml. The protein concentration of each fraction was calculated using the Bradford quantification method (Biorad). The glycine from the elution buffer was then removed first by pooling the fractions with the highest amount of protein, then adding to an Econo-Pac 10 DG column (Bio-Rad), allowing capture of the proteins and re-elution in a suitable buffer, sterile PBS. The proteins were eluted from the column in 1 ml of PBS; this elution step was repeated seven

more times. The protein concentrations in these eight fractions were again quantified using the Bradford method. The fractions with the most protein were pooled and concentrated to a final volume of approximately 0.8 – 1.0 ml using Vivaspin columns (Sartorius) and quantified again (Bio-Rad).

The affinity purified anti-RAD51 anti-serum was then used to probe the wildtype and *rad51* <sup>-/-</sup> protein extracts. For a direct comparison, two 10 % SDS-PAGE gels were equally loaded with protein extracts from wildtype bloodstream stage cells (wt) and *rad51* <sup>-/-</sup> cells (-/-). The separated proteins were then transferred to a nylon membrane. One membrane was probed with the unpurified rabbit anti-RAD51 anti-serum (at a dilution of 1:250) and the other was probed with the affinity purified anti-RAD51 anti-serum (at a dilution of 1:50) and both were detected with anti-rabbit IgG-HRP (Molecular Probes) anti-serum (at a 1:5000 dilution). Figure 5-27 shows these western blot analyses: Figure 5-27A shows the protein extracts probed with the unpurified anti-serum; in this blot, there was the same banding pattern seen in Figure 5-26 above, where there was a non-specific band of around 45 kDa, slightly larger than the RAD51 band. Figure 5-27B shows the protein extracts probed with the affinity purified anti-serum. The purified antibody did not bind to the non-specific band of approximately 45 kDa, but, surprisingly did bind to another protein of approximately 60 kDa. Removal of the non-specific binding to the protein running at approximately 45 kDa was important, as this could cause masking of the RAD51-anti-serum binding to the RAD51 paralogue protein, as they are all of a similar size, ranging from 40-55 kDa. Since bands of this size are not seen, it is likely that the anti-RAD51 anti-serum does not cross-react with the RAD51 paralogues.



**Figure 5-27: Western analysis showing the specificity of the unpurified and affinity purified anti-RAD51 anti-serum.**

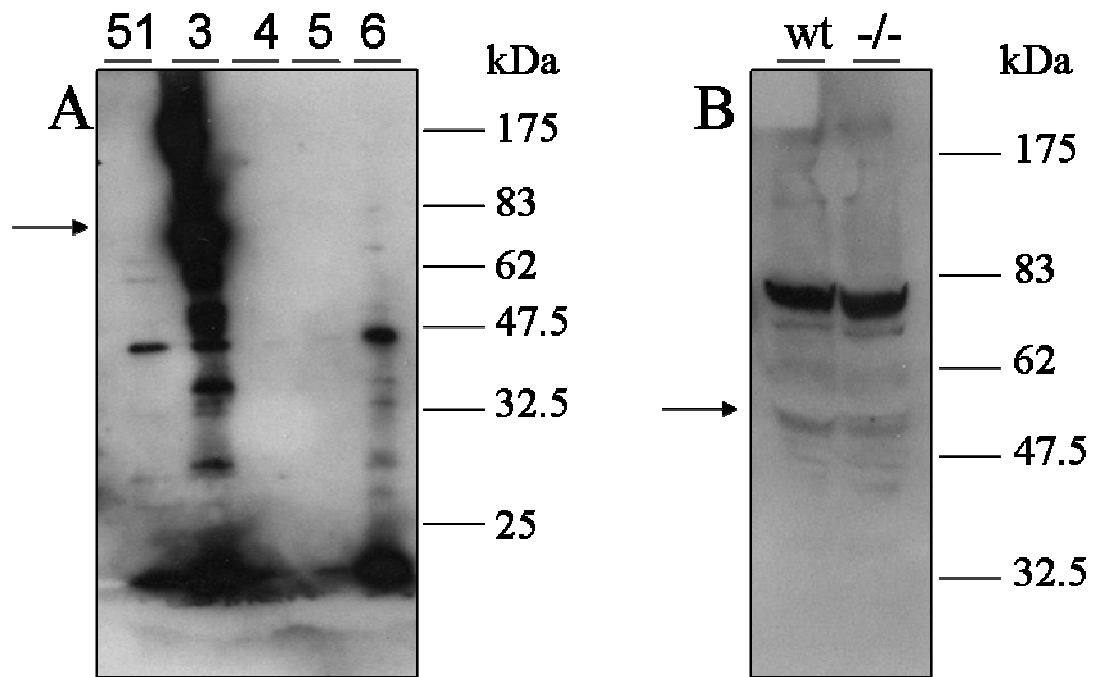
Proteins extracted from wildtype bloodstream form cells (wt) and *rad51* -/- (-/-) cells were separated on two 10 % SDS-PAGE gels and the proteins were then transferred to a nylon membrane. **Figure 5-27A:** shows the membrane probed with the unpurified anti-RAD51 anti-serum. **Figure 5-27B:** shows the membrane probed with the affinity purified anti-RAD51 anti-serum. Both figures show the predicted band of RAD51 at approximately 40 kDa, and the size markers are indicated for reference.

### 5.5.2 Anti-RAD51-3 anti-serum purification

The next anti-serum to be analysed was the sheep anti-RAD51-3 anti-serum. Firstly, the anti-serum was used to probe the GST-tagged RAD51 and RAD51 paralogue recombinant proteins. Approximately 2 ng.μl<sup>-1</sup> of each purified GST-recombinant protein was separated on a 10 % SDS-PAGE gel (see Figure 5-25B). The proteins were then transferred to a nylon membrane and probed firstly, with the sheep unpurified anti-RAD51-3 anti-serum (at a dilution of 1:1000) and secondly, detected with anti-sheep IgG-HRP (Santa Cruz) anti-serum (at a 1:5000 dilution; Figure 5-28A). The predicted size of the GST-RAD51-3 fusion is 80.8 kDa, and the black arrow in Figure 5-28A shows the predicted location of the GST-RAD51-3, amongst considerable cross-reaction. This analysis shows that anti-RAD51-3 anti-serum bound specifically to the GST-RAD51-3 recombinant protein, while no other bands of the predicted size for GST-RAD51, GST-RAD51-4, GST-RAD51-5 or GST-RAD51-6 were detected in these lanes.

Next, the specificity of the anti-RAD51-3 anti-serum was tested against whole cell protein extracts from *T. brucei* cells: wildtype BSF strain 427 protein extracts and BSF strain 427 *rad51-3* homozygous mutant (-/-) protein extracts. The extracted proteins were separated on a 10 % SDS-PAGE gel then transferred to a nylon membrane. The membrane was probed with the unpurified anti-RAD51-3 anti-serum (at a dilution of 1:1000) and detected

with anti-sheep IgG-HRP (Santa Cruz) anti-serum (at a 1:5000 dilution). Figure 5-28B does not show any distinct band of the predicted size of RAD51-3 (54.5 kDa) in the wildtype protein extract (wt), compared with the *rad51-3*  $-/-$  protein extract. The predominant reaction was to a non-specific cross-reactive band of ~80 kDa in both cell extracts, and so affinity purification of the anti-serum was required for further analysis.

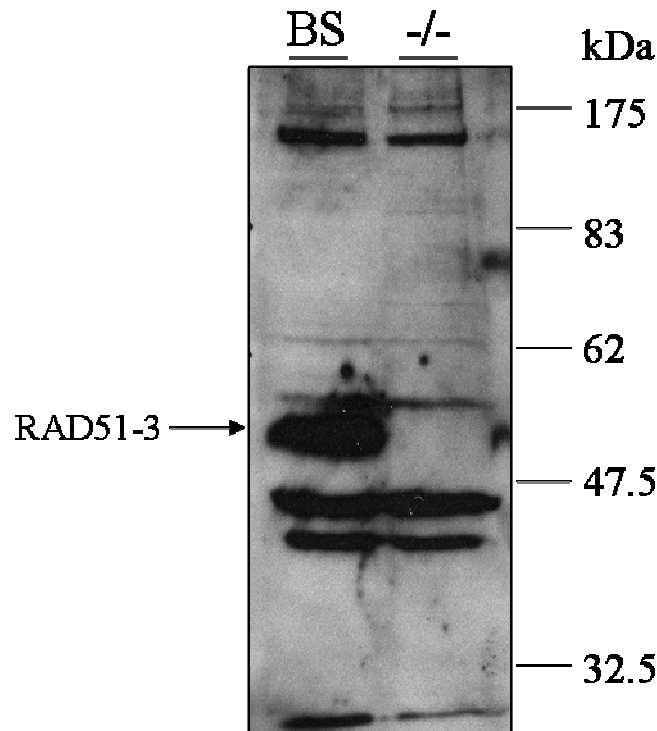


**Figure 5-28: Western analysis showing the specificity of the anti-RAD51-3 anti-serum**

**Figure 5-28A:** Western blot shows the separated GST-RAD51 and GST-RAD51 paralogue recombinant proteins probed with anti-RAD51-3 anti-serum. [Labelled: RAD51 (51), RAD51-3 (3), RAD51-4 (4), RAD51-5 (5), and RAD51-6 (6)]. The GST-tag adds 26.3 kDa to RAD51 and the RAD51 paralogues. The predicted sizes of the fusion proteins are GST-RAD51: 66.3 kDa, GST-RAD51-4: 72 kDa, GST-RAD51-5: 67.9 kDa and GST-RAD51-6: 76.7 kDa. The black arrow indicates the predicted size of the GST-RAD51-3 protein of 80.8 kDa. **Figure 5-28B:** Western blot shows protein extracts from the wildtype (wt) and *rad51-3* homozygous mutant ( $-/-$ ) *T. brucei* cells probed with the anti-RAD51-3 anti-serum. The black arrow indicates the predicted size of RAD51-3 (54.5 kDa). The size markers are indicated for both blots.

As a result of the poor specificity of the anti-RAD51-3 anti-serum, affinity purification was performed to improve the antibody's specificity. This was carried out in a similar manner to that described above for the anti-RAD51 anti-serum (see Section 5.5.1 and Materials and Methods, Section 2.7.2). The purified GST-RAD51-3 protein (described in Figure 5-25B) was coupled to beads which were then used as an affinity column. Approximately 5 mls of anti-RAD51-3 anti-serum were added to the GST-RAD51-3 affinity column. The column was washed thoroughly, and then the antibody was eluted from the beads. The fractions with the most protein were pooled and concentrated to approximately 800  $\mu$ l to 1 ml. The affinity purified anti-RAD51-3 anti-serum was then used to probe the wildtype and *rad51-3*  $-/-$  protein extracts. Figure 5-29 shows the affinity purified anti-RAD51-3 anti-serum (at a dilution of 1:10) used to probe protein extracts from wildtype BSF (BS) and

*rad51-3*  $-/-$  cells ( $-/-$ ); anti-serum binding was detected with anti-sheep IgG-HRP (Santa Cruz) anti-serum (at a 1:5000 dilution). Affinity purification greatly improved specificity to endogenous *T. brucei* RAD51-3 as there was a band of the predicted size of the protein (54.5 kDa) that was only seen in the wildtype extract and was absent in the *rad51-3*  $-/-$  mutant. At this anti-serum dilution, significant cross-reactivity was also seen.



**Figure 5-29: Western analysis showing the specificity of the affinity purified anti-RAD51-3 anti-serum.** Proteins extracted from wildtype bloodstream stage (BS) and *rad51-3*  $-/-$  mutant cells ( $-/-$ ) were separated on a 10 % SDS-PAGE gel. The proteins were transferred to a nylon membrane and then probed with the affinity purified anti-RAD51-3 anti-serum at a dilution of 1:10. The predicted size of RAD51-3 (54.5 kDa) and the size markers are indicated.

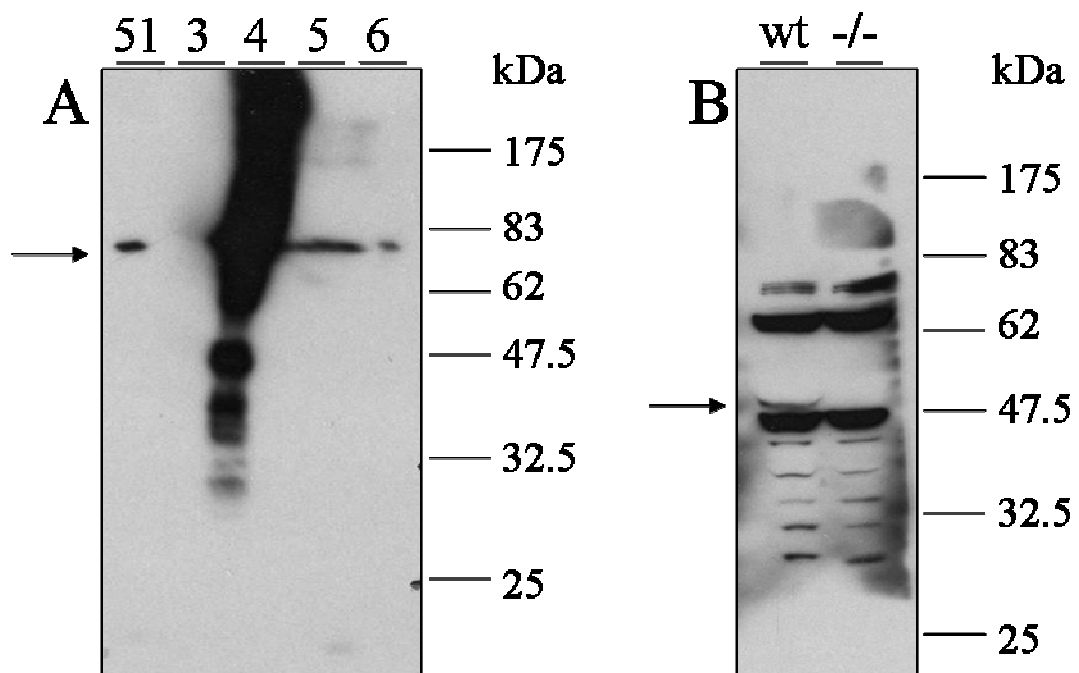
### 5.5.3 Anti-RAD51-4 anti-serum purification

Rabbit anti-RAD51-4 anti-serum was next analysed to test its specificity and see whether or not it bound to RAD51-4. To this end, the unpurified anti-serum was used to probe against the GST-tagged RAD51 and RAD51 paralogue recombinant proteins. As before, approximately  $2 \text{ ng.}\mu\text{l}^{-1}$  of each purified GST-recombinant protein was separated on a 10 % SDS-PAGE gel. The separated proteins were then transferred to a nylon membrane and probed with the unpurified rabbit anti-RAD51-4 anti-serum (at a dilution of 1:250), and detected with anti-rabbit IgG-HRP (Molecular Probes) anti-serum (at a 1:5000 dilution; Figure 5-30A). The predicted size of the GST-RAD51-4 fusion is 62 kDa, and in Figure 5-30A the black arrow indicates the predicted location of the GST-RAD51-4 protein. The western blot indicates that the anti-RAD51-4 anti-serum binds to the GST-RAD51-4 recombinant protein. It was also clear that the anti-serum bound to proteins of



approximately the same size as GST-RAD51-4 in the GST-RAD51, GST-RAD51-5 and GST-RAD51-6. These non-specific bands are approximately the same size in all the lanes and are not the predicted sizes of GST-RAD51, GST-RAD51-5 and GST-RAD51-6 (66.3 kDa, 67.9 kDa, and 76.7 kDa respectively), suggesting the anti-serum recognises an *E. coli*-derived protein of a similar size to the GST-RAD51-4 in these purifications.

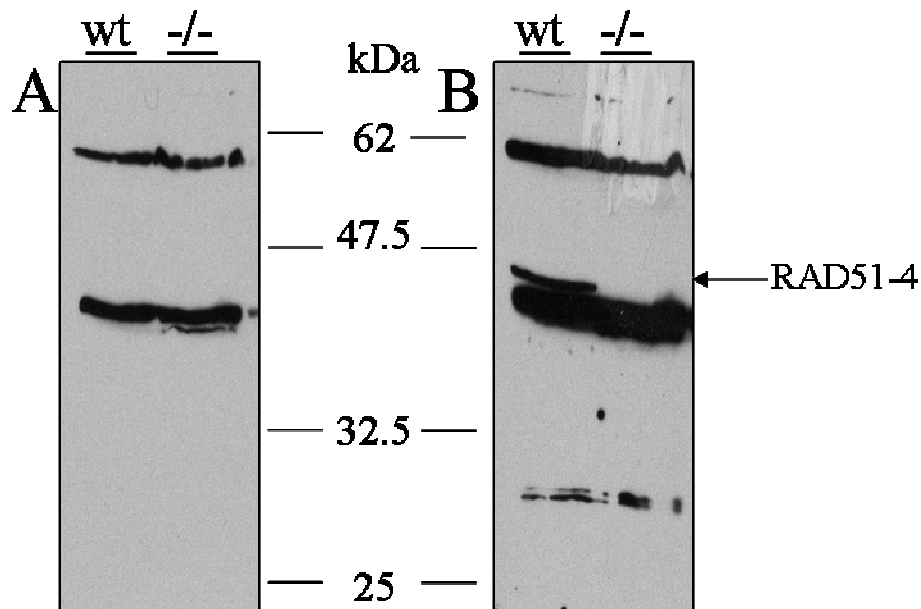
The specificity of rabbit anti-RAD51-4 anti-serum was next tested against whole cell protein extracts from *T. brucei* cells: wildtype BSF strain 427 protein extracts and BSF strain 3174.2 *rad51-4* homozygous mutant (-/-) protein extracts. The generation and confirmation of this *rad51-4* -/- mutant was described in Chapter 3, Section 3.2.3 of this thesis. The extracted proteins were separated on a 10 % SDS-PAGE gel then transferred to a nylon membrane. This was probed with the unpurified anti-RAD51-4 anti-serum at a dilution of 1:400 and detected with anti-rabbit IgG-HRP (Molecular Probes) anti-serum (at a 1:5000 dilution), see Figure 5-30B. A faint but distinct band of the predicted size of RAD51-4 (45.7 kDa) could be seen in the wildtype protein extract (wt) compared to the *rad51-4* -/- protein extract where it was absent. There was also considerable cross-reaction, but this suggests the anti-serum can recognise endogenous *T. brucei* RAD51-4.



**Figure 5-30: Western analysis showing the specificity of the anti-RAD51-4 anti-serum**

**Figure 5-30A:** Western blot shows the separated GST-RAD51 and GST-RAD51 paralogue recombinant proteins probed with anti-RAD51-4 anti-serum. [Labelled: RAD51 (51), RAD51-3 (3), RAD51-4 (4), RAD51-5 (5), and RAD51-6 (6)]. The GST tag adds 26.3 kDa to RAD51 and the RAD51 paralogues. The predicted sizes of the fusion proteins are GST-RAD51: 66.3 kDa, GST-RAD51-3: 80.8 kDa, GST-RAD51-5: 67.9 kDa and GST-RAD51-6: 76.7 kDa. The black arrow indicates the predicted size of the GST-RAD51-4 protein of 72 kDa. **Figure 5-30B:** Western blot shows the protein extract from wildtype (wt) and *rad51-4* homozygous mutant (-/-) *T. brucei* cells probed with the anti-RAD51-4 anti-serum. The black arrow indicates the predicted size of native RAD51-4, 45.7 kDa. The size markers are indicated for both blots.

Given the level of cross-reactive bands that were seen when the anti-RAD51-4 anti-serum was used to probe both the GST-tagged RAD51 and GST-tagged RAD51 paralogue recombinant proteins, as well as the wildtype and *rad51-4*  $-/-$  cell extracts, affinity purification of the anti-serum was required for further analysis. This was done using the GST-tagged RAD51-4 recombinant protein coupled to beads and used in the affinity column. Approximately 4 ml of anti-RAD51-4 anti-serum was added to the affinity column. Incubation for 2 hours at room temperature allowed the anti-RAD51-4 anti-serum to bind to the GST-RAD51-4 and this was then washed thoroughly to allow all unbound proteins to be discarded (Materials and Methods, Section 2.7.2). The purified anti-RAD51-4 anti-serum was then eluted from the GST-RAD51-4 beads in fractions of 0.5 ml. The protein concentration of each fraction was calculated and then the fractions with the most protein were pooled and concentrated. The unpurified and affinity purified anti-RAD51-4 anti-serum was then used to probe wildtype and *rad51-4*  $-/-$  protein extracts (Figure 5-31). The same amount of proteins extracted from wildtype (wt) and from bloodstream stage *rad51-4*  $-/-$  cells were separated on two 10 % SDS-PAGE gels, and the protein transferred to a nylon membrane. Figure 5-31B shows the membrane probed with the affinity purified anti-RAD51-4 antibody at a dilution of 1:100 and detected with anti-rabbit IgG-HRP (Molecular Probes) anti-serum (at a 1:5000 dilution). A distinct band of the predicted size of RAD51-4 (45.7 kDa) was seen in the wildtype protein extract but was absent in the *rad51-4*  $-/-$  cell extract. Comparisons of the affinity purified anti-RAD51-4 anti-serum with the unpurified anti-serum show that the detection of endogenous RAD51-4 is much more distinct in the former (compare Figure 5-31B and Figure 5-31A). However, there were more non-specific bands detected by the affinity purified anti-RAD51-4 anti-serum, and the pronounced band just smaller than 45 kDa remained.

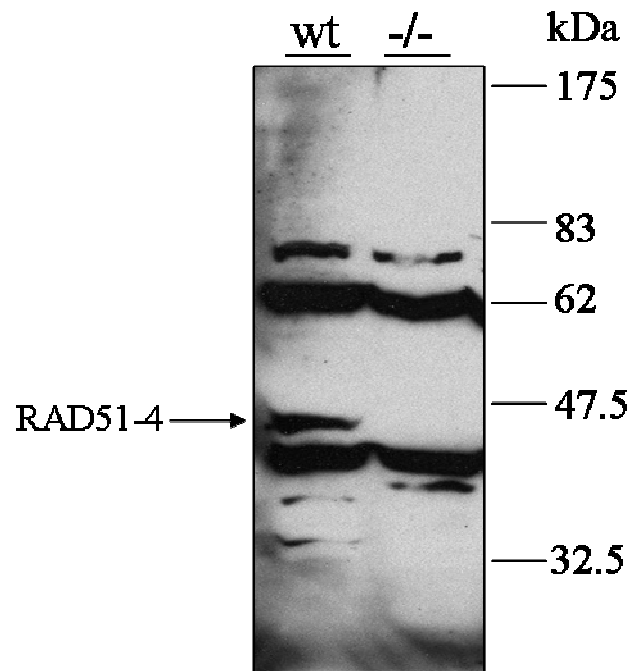


**Figure 5-31: Western analysis showing the specificity of the affinity purified anti-RAD51-4 anti-serum.** Whole cell proteins extracted from wildtype bloodstream stage (wt) and *rad51-4*  $-/-$  mutant cells were separated on two 10 % SDS-PAGE gels, and then transferred to a nylon membrane. **Figure 5-31A:** shows the membrane probed with the unpurified anti-RAD51-4 anti-serum at a dilution of 1:500. **Figure 5-31B:** shows the membrane probed with the affinity purified anti-RAD51-4 antibody at a dilution of 1:100. The predicted size of RAD51-4 (45.7 kDa) is indicated by the black arrow. The size markers are indicated for both blots.

In an attempt to further purify the anti-serum and remove the detection of non-specific proteins, another method was adopted. This method uses an acetone extract of *T. brucei* *rad51-4*  $-/-$  cells, which do not to express the target antigen, *T. brucei* RAD51-4 (see Material and Methods, Section 2.7.3 for the protocol). As there is no RAD51-4 in the protein in the extract from the *rad51-4*  $-/-$  cells, addition of the anti-RAD51-4 anti-serum will result in non-specific antibodies binding to their target antigens in the total extracted protein while the anti-RAD51-4 antibodies remaining unbound and free in the supernatant. A total of  $2 \times 10^9$  *rad51-4*  $-/-$  cells were used to purify 1 ml of anti-RAD51-4 anti-serum. Briefly, 500 ml of *rad51-4*  $-/-$  at a density of  $2 \times 10^6$  cells per ml were pelleted, washed in PBS and then washed twice with cold 100 % acetone. In between each wash the cells were incubated on ice for 10 minutes. The final pellet was spread on filter paper and allowed to completely dry. After drying, the protein was placed in an eppendorf tube and the unpurified anti-sera was added to it. This solution was gently mixed and incubated at 4 °C for 30 minutes, and then spun for 30 minutes. The resulting supernatant contains the purified antibodies.

To test if the acetone extraction method purified the anti-serum more effectively than the affinity purification method, the *T. brucei* wildtype and *rad51-4*  $-/-$  protein extracts were separated on a 10 % SDS-PAGE gel and transferred to a nylon membrane. Figure 5-32 shows the membrane probed with purified anti-serum at a dilution of 1:500. This was then

detected using anti-rabbit IgG-HRP (Molecular Probes) anti-serum (at a 1:5000 dilution). Western blot analysis showed that there was a band present in the wildtype extract which was absent in the *rad51-4*  $-/-$  cell extract and is the predicted size of RAD51-4 (45.7 kDa). This purification method did not remove the non-specific cross-reactive bands and was not more efficient than affinity purification. As a result the affinity purified anti-RAD51-4 anti-serum was used for further analyses.



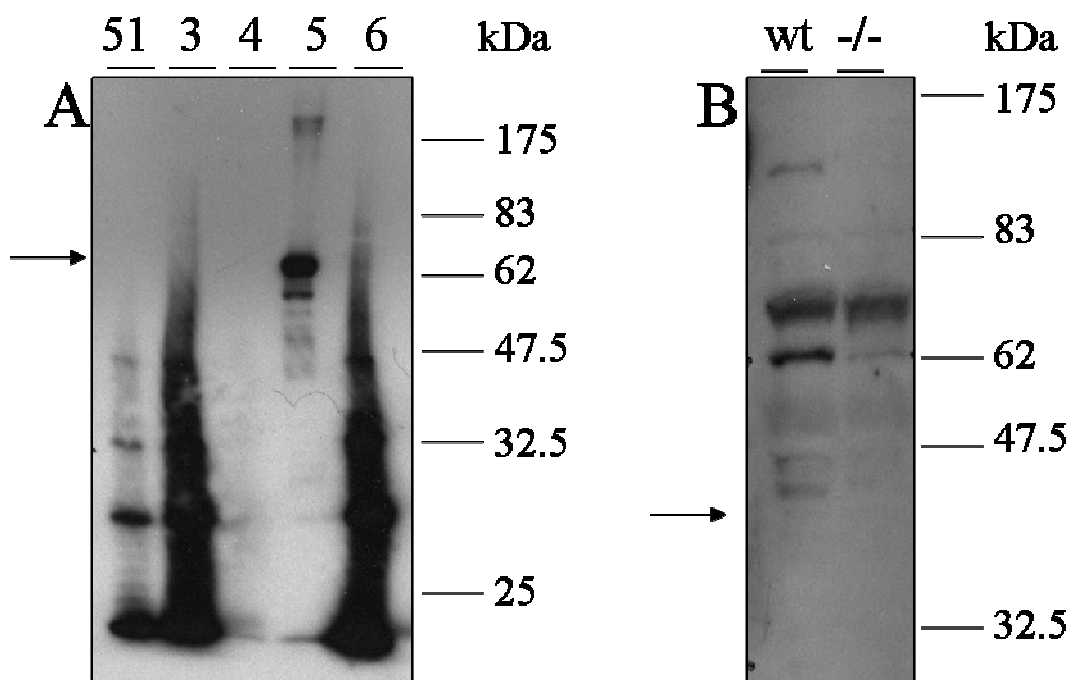
**Figure 5-32: Western analysis showing the specificity of the purified anti-RAD51-4 anti-serum using the acetone extraction method of total proteins from cells not expressing RAD51-4.**

Whole cell proteins extracted from wildtype bloodstream stage cells (wt) and *rad51-4*  $-/-$  mutant cells ( $-/-$ ) were separated on a 10 % SDS-PAGE gel, and transferred to a nylon membrane. The membrane was probed with the purified anti-RAD51-4 anti-serum at a dilution of 1:500. This anti-serum was purified using the acetone extraction method. A band of the predicted size of RAD51-4 (45.7 kDa) is indicated with the black arrow. The size markers are indicated for the blot.

#### 5.5.4 Anti-RAD51-5 anti-serum purification

The next anti-serum to be analysed was the chicken anti-RAD51-5 anti-serum. Firstly, the polyclonal antibodies were purified from the chicken egg yolk produced. This was carried out with the use of a Chicken IgY purification kit from Pierce (as per the manufacturer's instructions and see Materials and Methods, Section 2.7.1). Once the polyclonal anti-serum had been extracted from the egg, it was used to probe the GST-tagged RAD51 and RAD51 paralogue recombinant proteins (see Figure 5-25B). Each purified GST-recombinant protein was separated on a 10 % SDS-PAGE gel and western blot analysis was carried out. The proteins were transferred to a nylon membrane and probed with the chicken anti-RAD51-5 anti-serum (at a dilution of 1:500) and detected with anti-chicken IgG-HRP (Jackson ImmunoResearch) anti-serum (at a 1:7500 dilution), see Figure 5-33A. The

predicted size of the GST-RAD51-5 fusion is 67.9 kDa and a band of the correct size was seen in the lane loaded with GST-RAD51-5. The band was absent in the other lanes containing GST-RAD51, GST-RAD51-3, GST-RAD51-4 and GST-RAD51-6, indicating that the anti-serum was specific for RAD51-5. Figure 5-33B shows the anti-RAD51-5 anti-serum used to probe the *T. brucei* wildtype cell protein extract alongside *T. brucei rad51-5*  $-/-$  mutant cell protein extract (Proudfoot & McCulloch, 2005a). Again, the membrane was probed with the anti-RAD51-5 anti-serum (at a dilution of 1:500) and detected with anti-chicken IgG-HRP (Jackson ImmunoResearch) anti-serum (at a 1:7500 dilution; Figure 5-33B). The western blot analysis does not show any specific band of the predicted size of RAD51-5 (41.6 kDa) in the wildtype (wt) compared with the *rad51-5*  $-/-$  cell protein extracts (-/-). The anti-RAD51-5 anti-serum detected a number of non-specific cross-reactive bands in both cell extracts, and therefore affinity purification was required for further analysis.



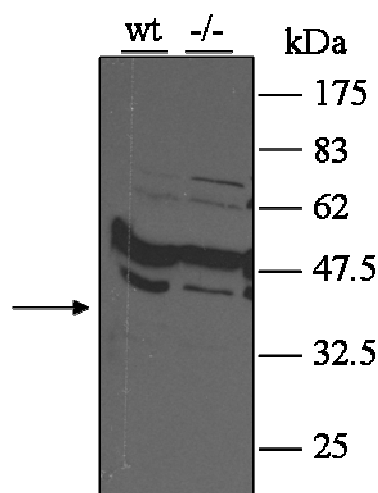
**Figure 5-33: Western analysis showing the specificity of the anti-RAD51-5 antibodies purified from chicken eggs.**

**Figure 5-33A:** Western blot shows the separated GST-RAD51 and GST-RAD51 paralogue recombinant proteins probed with anti-RAD51-5 anti-serum. [Labelled: RAD51 (51), RAD51-3 (3), RAD51-4 (4), RAD51-5 (5), and RAD51-6 (6)]. The GST-tag adds 26.3 kDa to RAD51 and the RAD51 paralogs as a result the predicted sizes of the fusion proteins are GST-RAD51: 66.3 kDa, GST-RAD51-3: 80.8 kDa, GST-RAD51-4: 72 kDa, and GST-RAD51-6: 76.7 kDa. The black arrow indicates the predicted size of the GST-RAD51-5 protein of 67.9 kDa. **Figure 5-33B:** Western blot shows proteins extracted from wildtype (wt) and *rad51-5* homozygous mutant (-/-) *T. brucei* cells probed with the anti-RAD51-5 anti-serum. The black arrow indicates the predicted size of RAD51-5, 41.6 kDa. The size markers are indicated for both blots.

Affinity purification was performed using the same method as described above (Sections 5.5.1, 5.5.2, and 5.5.3), and used GST-RAD51-5 coupled beads to bind to the anti-RAD51-

5 anti-serum. Non-specific antibodies in the anti-serum were washed away in the flow-through. The anti-RAD51-5 antibodies were then eluted off the beads, quantified and concentrated.

Unfortunately, the affinity purification of the anti-RAD51-5 anti-serum was not effective. The anti-serum was used to probe wildtype and *rad51-5*  $-/-$  protein extracts, but no specific bands were detected of the predicted size of RAD51-5 (41.6 kDa) in the wildtype compared with *rad51-5*  $-/-$  mutant cell extracts, when probed with 1:50 dilution of affinity purified anti-RAD51-5. This was detected with anti-chicken IgG-HRP (Jackson ImmunoResearch) anti-serum (at a 1:7500 dilution; Figure 5-34).



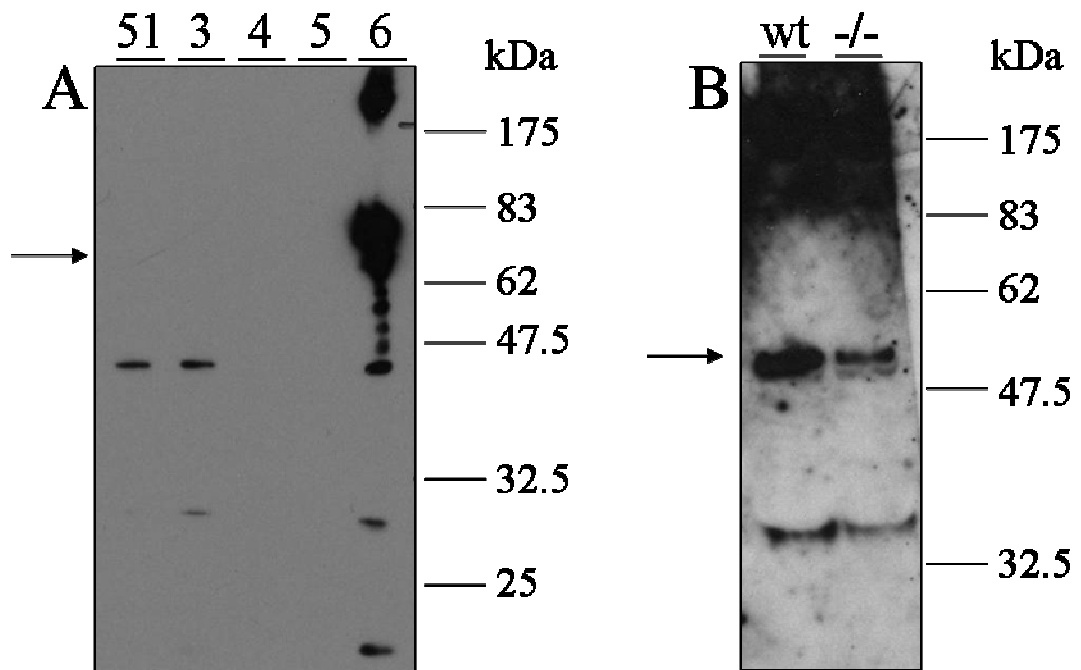
**Figure 5-34: Western analysis showing the specificity of the affinity purified anti-RAD51-5 anti-serum.** Proteins extracted from wildtype (wt) and *rad51-5*  $-/-$  mutant cells ( $-/-$ ) were separated on a 10 % SDS-PAGE gel. The proteins were transferred to a nylon membrane and then probed with the affinity purified anti-RAD51-5 anti-serum at a dilution of 1:50. The predicted size of RAD51-5 (41.6 kDa) is indicated by the black arrow and the size markers are also shown.

If more time was available, one possible way of overcoming this would be to try a range of different egg yolks. The egg yolk used here was taken from an egg produced late after the injection of the recombinant protein (Day 26) and so extracting the antibodies from an earlier egg yolk may result in more specific anti-RAD51-5 antibodies. Time constraints did not allow this to be done. Instead, to overcome the lack of reactivity of the anti-serum, PTP-tagging of the RAD51-5 was performed to allow protein affinity purification: this is described in Section 5.7 of this chapter.

### 5.5.5 Anti-RAD51-6 anti-serum purification

The final anti-serum to be analysed was rat anti-RAD51-6 anti-serum. Like the other anti-sera, it was important to test its specificity and see whether or not it bound specifically to *T. brucei* RAD51-6. To this end, the unpurified anti-RAD51-6 anti-serum was used to probe the GST-tagged RAD51 and RAD51 paralogue recombinant proteins. As before, each purified GST-recombinant protein was separated on a 10 % SDS-PAGE gel, and the proteins were transferred to a nylon membrane. The membrane was probed with the unpurified anti-RAD51-6 anti-serum (at a dilution of 1:500) and detected with anti-rat IgG-HRP (Sigma) anti-serum (at a dilution of 1:10,000; Figure 5-35A). The predicted size of the GST-RAD51-6 fusion is 76.7 kDa and a distinct reactivity of the anti-RAD51-6 anti-serum to a protein of this size was seen (Figure 5-35A, Lane 6), while there was no evidence of cross-reactivity against GST-RAD51 and the other GST-RAD51 paralogue recombinant proteins.

Figure 5-35B shows the anti-RAD51-6 anti-serum used to probe *T. brucei* wildtype and *T. brucei rad51-6 -/-* protein extracts. The generation and confirmation of the BSF strain 427 *rad51-6 -/-* was described in this thesis (Chapter 3, Section 3.5.3). Both protein extracts were separated on a 10 % SDS-PAGE gel and transferred to a nylon membrane. The membrane was probed with the anti-RAD51-6 anti-serum (at a dilution of 1:250) and detected with anti-rat IgG-HRP (Sigma) anti-serum (at a dilution of 1:10,000; Figure 5-35B). Specific bands of the predicted size of RAD51-6 (50.4 kDa) were not seen in the wildtype extract compared with the *rad51-6 -/-* protein extract. All bands appear instead to be non-specific, cross-reactive bands in both cell extracts (approximate sizes of 52 kDa and 35 kDa) and affinity purification of anti-serum was required for further use of the antibody.

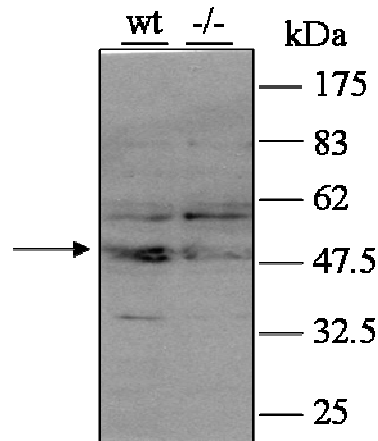


**Figure 5-35: Western analysis showing the specificity of the anti-RAD51-6 anti-serum**

**Figure 5-35A:** Western blot shows the separated GST-RAD51 and GST-RAD51 paralogue recombinant proteins probed with anti-RAD51-6 anti-serum. [Labelled: RAD51 (51), RAD51-3 (3), RAD51-4 (4), RAD51-5 (5), and RAD51-6 (6)]. This GST tag fusion added 26.3 kDa to RAD51 and the RAD51 paralogues. The predicted sizes of the fusion proteins are GST-RAD51: 66.3 kDa, GST-RAD51-3: 80.8 kDa, GST-RAD51-4: 72 kDa and GST-RAD51-5: 67.9 kDa. The black arrow indicates the predicted size of the GST-RAD51-6 protein of 76.7 kDa. **Figure 5-35B:** Western blot shows the separated wildtype (wt) and *rad51-6* homozygous mutant (-/-) protein extracts. The black arrow indicates the predicted size of RAD51-6, 50.4 kDa. The size markers are indicated for both blots.

Affinity purification was completed using the same method as described above, (Section 5.5.1, 5.5.2, 5.5.3, and 5.5.4), and used GST-RAD51-6 coupled beads to bind to the anti-RAD51-6 antibodies. This allowed the non-specific antibodies in the anti-serum to be washed away in the flow-through. The anti-RAD51-6 antibodies were then eluted off the beads, quantified and concentrated. Similarly to the anti-RAD51-5 affinity purification, the affinity purification of anti-RAD51-6 antibodies was not effective. Wildtype and *rad51-6* -/- cell protein extracts were separated on a 10 % SDS-PAGE gel, transferred to a nylon membrane and probed with the affinity purified anti-RAD51-6 anti-serum (at a dilution of 1:10). Anti-rat IgG-HRP (Sigma) anti-serum (at a dilution of 1:10,000) was used to detect the primary antibody. This western blot analysis detected no distinct bands in the wildtype extracts compared to *rad51-6* -/- mutant at the predicted size of RAD51-6 (~ 50 kDa) (Figure 5-36).





**Figure 5-36: Western analysis showing the specificity of the affinity purified anti-RAD51-6 anti-serum.** Proteins extracted from wildtype bloodstream stage (wt) and *rad51-6*  $-/-$  mutant cells ( $-/-$ ) were separated on a 10 % SDS-PAGE gel. The proteins were transferred to a nylon membrane and then probed with the affinity purified anti-RAD51-6 anti-serum at a dilution of 1:10. The predicted size of RAD51-6 (50.4 kDa) is indicated by the black arrow and the size markers are also shown.

To overcome the fact that the anti-RAD51-6 anti-serum could not be used for further analyses such as co-immunoprecipitation, a PTP-tagged RAD51-6 was generated, and is described in Section 5.7.

## 5.6 Co-immunoprecipitation of the RAD51 paralogues

### 5.6.1 Principles of co-immunoprecipitation

Co-immunoprecipitation (Co-IP) is a method in which a target protein, or other antigen, is precipitated from a cell extract using specific antibodies that recognise the target protein or antigen. This technique involves immobilising the specific antiserum, and incubating it with cell extract. This allows binding of the protein or antigen from the cell extract solution to the specific antibodies. As a result the anti-serum not only immobilises the target protein or antigen but can also isolate other proteins with which it interacts, as long as the complex is maintained in the cell extract. This method involves coupling the anti-serum to Protein A, Protein G beads or gel matrix, and allowing the bound anti-serum to bind to its target antigen along with any protein or protein complexes attached to the target antigen.

For the analysis, approximately 1-2 mg of extracted protein was prepared from wildtype cells and the *rad51* paralogue homozygous mutant cells. A gentle method of whole cell protein extraction was performed to avoid disrupting the protein interactions; first the nuclear content was extracted, this was then pooled with the proteins extracted from the

non-nuclear material (see Materials and Methods, Section 2.6.2.2 for protocol). Briefly, the cells were grown to a density of  $2\text{--}3 \times 10^6$  per ml, pelleted, washed twice with PBS and frozen quickly using dry-ice and stored at  $-80^\circ\text{C}$  until required. The frozen pellets were resuspended in hypotonic lysis buffer, incubated on ice for 20 minutes, vortexed for 30 seconds and pelleted at  $800 \times g$  for 2 minutes to collect the nuclei. The supernatant was retained and stored on ice. The pelleted nuclei were resuspended in nuclear extract buffer and incubated on ice for 20 minutes. The retained supernatant was added back to the reaction and the cell debris was removed by a centrifugation at  $20,000 \times g$  for 20 minutes at  $4^\circ\text{C}$ . This supernatant contained protein extracted from the entire cell and was quantified and used in the co-IP analyses.

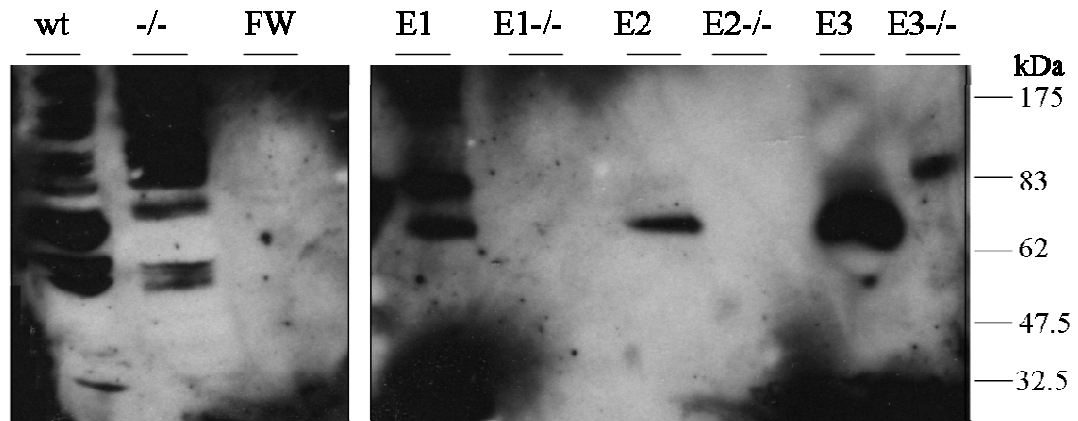
The Profound™ co-immunoprecipitation kit (Pierce) was used for the co-IP analyses here. The analysis was performed in a 1.5 ml eppendorf tube with the advantage that all spins can be completed in a table top centrifuge. This kit links the antibody using an “antibody coupling gel” to a gel matrix, with the retention of the anti-serum after the final elution step. This aids the analysis, as other methods that result in elution of the anti-serum along with the proteins may have the proteins of interest masked by the light or heavy chains of the antibody. The protein extract was added to the antibody-coupled-resin, incubated for 2 hours at room temperature, and then washed thoroughly with the wash buffers provided. The protein bound to the anti-serum was then eluted off the antibody-coupled-resin in fractions. These fractions were then analysed by western blot.

Section 5.5 above describes the affinity purification of the RAD51 and RAD51 paralogue anti-sera. It was concluded that only the affinity purified anti-RAD51-3 and anti-RAD51-4 anti-sera could be used in co-immunoprecipitation (Co-IP) analyses. The lack of specific reaction by the anti-RAD51-5 and anti-RAD51-6 anti-sera prevented their use.

### **5.6.2 Co-immunoprecipitation using the affinity purified anti-RAD51-3 anti-serum**

To perform the co-IP analyses,  $400\text{ }\mu\text{g}$  of the affinity purified anti-RAD51-3 anti-serum was added to the antibody-coupled-resin. This was done for two independent tests. For one analysis, approximately 1.5 mg of protein extract was prepared from wildtype bloodstream stage *T. brucei* Lister 427 cells and from *rad51-3*  $-/-$  cells. The co-IP analysis was conducted using the manufacturer’s instructions (Pierce) and is described in the Materials and Methods, Section 2.11. Figure 5-37 shows the western blot of the input protein extract from wildtype and *rad51-3*  $-/-$  cells, the final wash from the wildtype cell extract, and the

three eluted fractions (E1, E2, E3). The membrane was probed with sheep anti-RAD51-3 anti-serum (at a dilution of 1:10) and detected using anti-sheep IgG-HRP (1:5000). The Pierce West Femto was used in these experiments, which is a maximum sensitivity chemiluminescent substrate; as a result there is an increase in the background of non-specific binding. There is a clear band in the wildtype of 50.4 kDa in size, compared to the *rad51-3*  $-/-$  mutant.

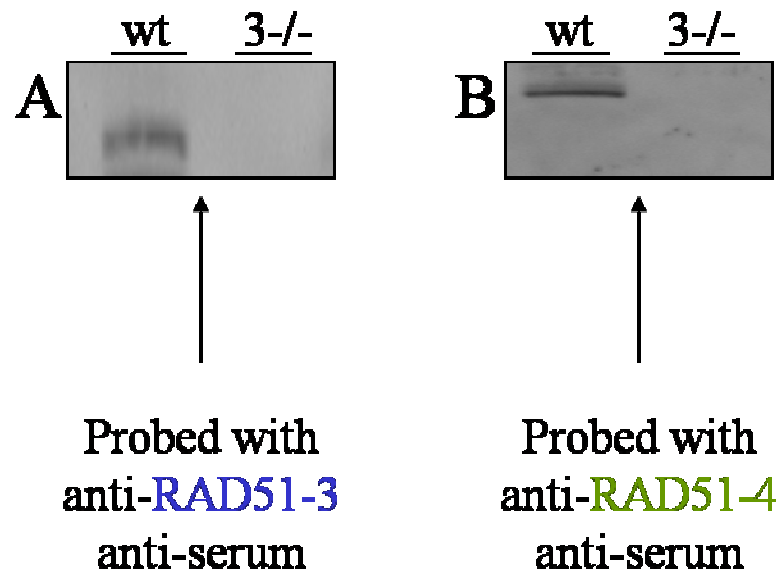


**Figure 5-37: Western analysis of the eluates following the co-immunoprecipitation using affinity purified anti-RAD51-3 anti-serum.**

Affinity purified RAD51-3 anti-serum was used to perform a co-IP from wildtype (wt) and *rad51-3*  $-/-$  ( $-/-$ ) cells. The input from the wildtype and *rad51-3*  $-/-$  mutant cells, the final wash of the beads containing the wildtype cell extract (FW) and the eluted fractions from the co-IP were separated on a 10 % SDS-PAGE gel and were then transferred to a nylon membrane. This was probed with sheep anti-RAD51-3 anti-serum. E1, E2 and E3 are the eluates of the co-IP from the wildtype cells. E1 $-/-$ , E2 $-/-$  and E3 $-/-$  are the eluates of the co-IP from the *rad51-3*  $-/-$  cells.

As there was a clean band of the correct size in the second fraction, the rest of the eluates from the co-IP of the wildtype and *rad51-3*  $-/-$  cells were loaded onto a 10 % SDS-PAGE gel in triplicate (~30 % of remaining second fraction). The proteins were then transferred to a nylon membrane, and this was cut into three. Each membrane was probed with the affinity purified rabbit anti-RAD51 anti-serum (at a dilution of 1:10), the affinity purified sheep anti-RAD51-3 anti-serum (at dilutions of 1:50) and the affinity purified rabbit anti-RAD51-4 anti-serum (at a dilution of 1:1000). These were detected using IgG-HRP antisera for each species, anti-rabbit (1:1000) or anti-sheep (1:5000). Figure 5-38 shows the results of this using the second eluted fraction from the co-IP. Figure 5-38A shows these fractions probed with anti-RAD51-3 anti-serum, demonstrating that the RAD51-3 was present in the eluate from the co-IP of wildtype but absent from the *rad51-3*  $-/-$  mutant. When the eluate was probed with anti-RAD51-4 anti-serum (Figure 5-38B) it was shown that RAD51-4 was also present from the co-IP from the wildtype cell extract, but not from the *rad51-3*  $-/-$  mutant cell extract. Anti-RAD51 anti-serum was used to probe the fraction, RAD51 was not detected in the elutant from the co-IP of either wildtype cell extract or

*rad51-3*  $-/-$ , and therefore indicating RAD51 interaction with RAD51-3 was not observed (data not shown). The analysis confirms that in the eluted fraction, RAD51-3 and RAD51-4 were both present indicating that these proteins interact in *T. brucei* while RAD51 was absent.



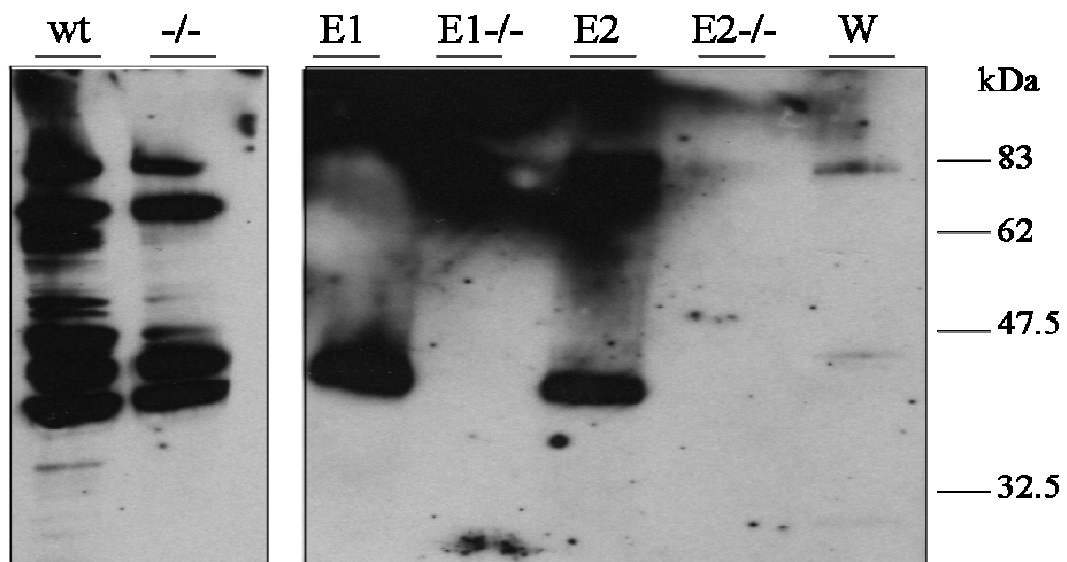
**Figure 5-38: Western analysis following the co-immunoprecipitation using affinity purified anti-RAD51-3 anti-serum.**

Affinity purified RAD51-3 anti-serum was used to perform a co-IP from wildtype and *rad51-3*  $-/-$  *T. brucei* bloodstream stage cells. Eluted fractions from the co-IP were separated on a 10 % SDS-PAGE gel and were then transferred to a nylon membrane. **Figure 5-38A:** Western blot of one fraction from the co-IP from wildtype (wt) and *rad51-3*  $-/-$  (3-/-) cells when probed with anti-RAD51-3 anti-serum. **Figure 5-38B:** Western blot of one fraction from co-IP of wildtype (wt) and *rad51-3*  $-/-$  (3-/-) cells when probed with the affinity purified anti-RAD51-4 anti-serum.

### 5.6.3 Co-immunoprecipitation using the affinity purified anti-RAD51-4 anti-serum

As the affinity purified anti-RAD51-4 anti-serum was specific to *T. brucei* RAD51-4 and had only a few cross-reactive bands, co-IP analyses was carried out using the ProFound™ co-immunoprecipitation kit (Pierce) as before (Materials and Methods, Section 2.11). Briefly, a total of 350 µg of the affinity purified anti-RAD51-4 anti-serum was added to the antibody-coupled-resin. Again, this was done in duplicate for two independent tests. Approximately 1.5 mg of protein extract from wildtype cells was added to the antibody-coupled-resin. The same amount of extracted protein from the *rad51-4*  $-/-$  mutant cells was added to the other antibody-coupled-resin. The generation and confirmation of *rad51-4*  $-/-$  mutant was completed and described in this thesis (Chapter 3, Section 3.2.3).

Figure 5-39 shows the input protein extract from wildtype (wt) and *rad51-4*  $-/-$  ( $-/-$ ) cells, the final wash (W) from the wildtype cell extract, and two eluted fractions (E1, and E2) were loaded onto a 10 % SDS-PAGE gel. The proteins were then transferred to a nylon membrane and probed with rabbit anti-RAD51-4 anti-serum (at a dilution of 1:500). These were detected using anti-rabbit IgG-HRP (1:5000). The Pierce West Femto, a maximum sensitivity chemiluminescent substrate, was used in these experiments. Use of this chemiluminescent substrate increases the non-specific binding in the wildtype and homozygous cell extracts, as well as increasing the background at the top of the membrane. As a result there is not a clear input band of 40.5 kDa in the wildtype compared to the *rad51-4*  $-/-$  cell extract because of the non-specific chemiluminescence. Despite this, there is a distinct band in the eluted fractions of the co-IP from the wildtype which is the correct size of 40.5 kDa and is not present in the co-IP from the *rad51-4*  $-/-$  cells.

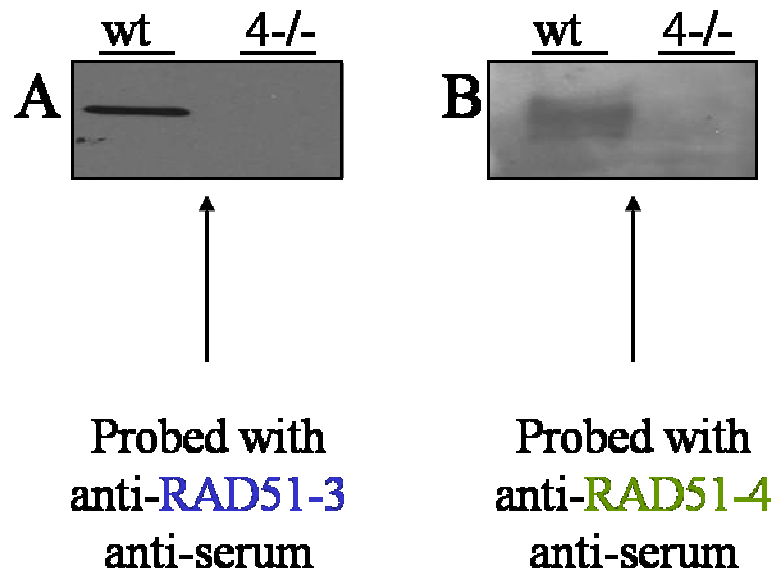


**Figure 5-39: Western analysis of the eluates following the co-immunoprecipitation using affinity purified anti-RAD51-4 anti-serum.**

Affinity purified RAD51-4 anti-serum was used to perform a co-IP from wildtype (wt) and *rad51-4*  $-/-$  ( $-/-$ ) cells. The input from the wildtype and *rad51-4*  $-/-$  mutant cells, the final wash of the beads containing the wildtype cell extract (W) and the eluted fractions from the co-IP were separated on a 10 % SDS-PAGE gel and were then transferred to a nylon membrane. This was then probed with rabbit anti-RAD51-4 anti-serum. E1 and E2 are the eluates of the co-IP from the wildtype cells. E1-/- and E2-/- are the eluates of the co-IP from the *rad51-4*  $-/-$  cells.

Using the affinity purified anti-RAD51-4 anti-sera, equal amounts of the first eluted fraction from the co-IP of the wildtype and *rad51-4*  $-/-$  cells were loaded in triplicate onto a 10 % SDS-PAGE gel with each separated by a size marker. The proteins were then transferred to a nylon membrane and the membrane cut into three parts. Each part of the membrane was probed with a different affinity purified anti-sera: rabbit anti-RAD51 anti-sera, sheep anti-RAD51-3 anti-sera and rabbit anti-RAD51-4 anti-sera (at a dilution of 1:10, 1:50, 1:1000 respectively) as described above, and were then detected using the IgG-

HRP anti-sera for each species, anti-rabbit and anti-sheep (at a dilution of 1:1000 and 1:5000 respectively). Figure 5-40A shows the result of the co-IP analysis when the eluted fraction was probed with the affinity purified anti-RAD51-3 antibody. This demonstrates that when the anti-RAD51-4 anti-serum was used to precipitate RAD51-4, RAD51-3 was also present in the eluate indicating that RAD51-4 interacts with RAD51-3. Figure 5-40B shows the same fractions probed with anti-RAD51-4 anti-serum, demonstrating that the RAD51-4 was present in the eluate from the co-IP wildtype cells while absent from the *rad51-4*  $-/-$  mutant. Again, when the eluate from the co-IP of either wildtype cell extract or *rad51-4*  $-/-$  was probed with anti-RAD51 anti-serum, no detectable bands were seen (data not shown), and indicate that RAD51 was absent from the complex. These results confirm that in the eluted fraction from the wildtype, RAD51-3 and RAD51-4 were both present.



**Figure 5-40: Western analysis following the co-immunoprecipitation using affinity purified anti-RAD51-4 anti-serum.**

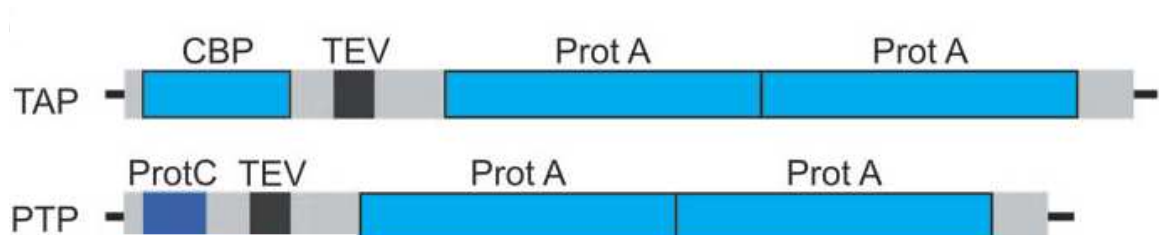
Affinity purified RAD51-4 anti-serum was used to perform a co-IP from wildtype and *rad51-4*  $-/-$  *T. brucei* bloodstream form cells. Eluted fractions were separated on a 10 % SDS-PAGE gel and were then transferred to a nylon membrane. **Figure 5-40A:** Western blot of one fraction from co-IP of wildtype (wt) and *rad51-4*  $-/-$  (4-/-) cells when probed with anti-RAD51-3 anti-serum. **Figure 5-40B:** Western blot of one fraction from co-IP of wildtype (wt) and *rad51-4*  $-/-$  (4-/-) cells when probed with the affinity purified anti-RAD51-4 anti-serum.

## 5.7 PTP-tag purification of protein complexes

### 5.7.1 Principles of PTP-tag purifications

The PTP-tag purification is a novel way of protein purification using a modified Tandem Affinity Purification (TAP) tag (Schimanski *et al.*, 2005b). The TAP-tag contains two immunoglobulin (IgG) binding domains of *Staphylococcus aureus* protein A (Prot A) and a

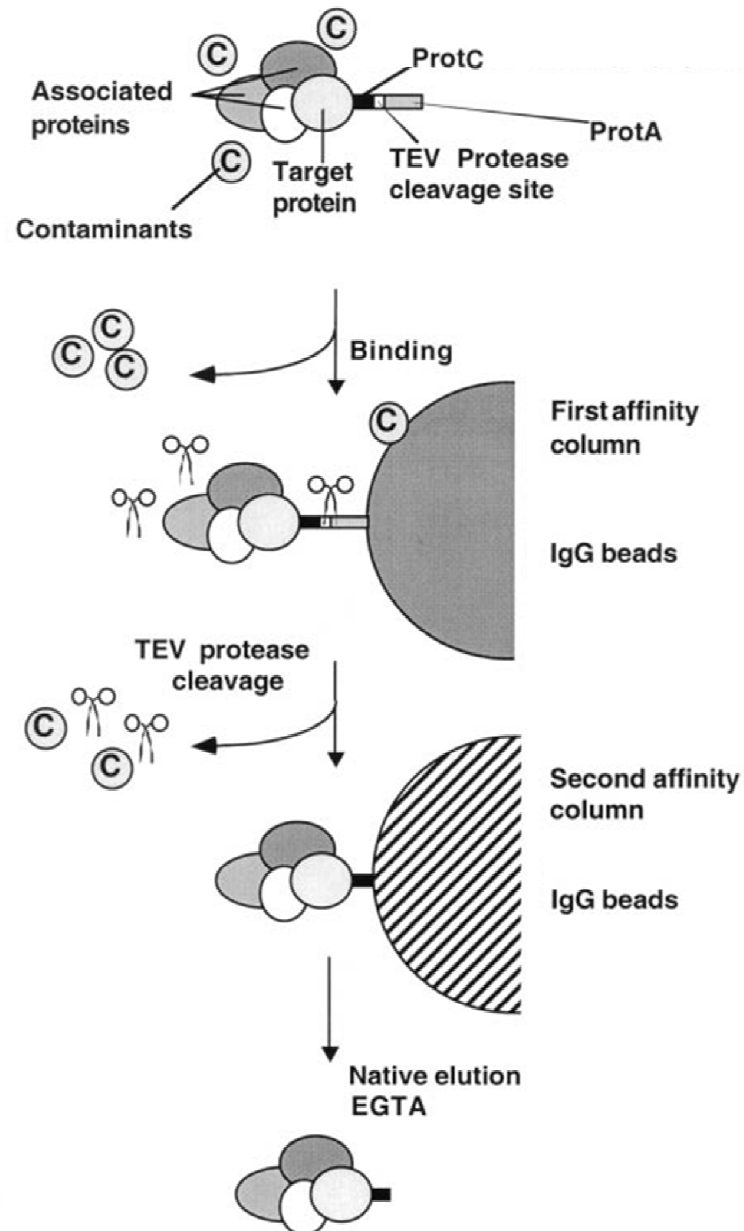
calmodulin binding peptide (CBP) separated by a tobacco etch virus (TEV) protease site (Kelly *et al.*, 2007; Puig *et al.*, 2001; Figure 5-41). The TAP-tag method allows purification of the protein of interest and associate proteins by fusion of the protein to the tag, expression of the tagged protein in the required cell line and purification of the protein or proteins in non-denaturing conditions. The Gunzl lab designed the PTP-tag by replacing the calmodulin binding peptide with the Protein C epitope (Prot C), see Figure 5-39. The PTP-tag contains the immunoglobulin (IgG) binding domain of *Staphylococcus aureus* protein A and the human Protein C separated by a tobacco etch virus (TEV) protease site (Schimanski *et al.*, 2005a).



**Figure 5-41: Schematic diagram showing the TAP-tag and PTP-tag.**

The Protein A (Prot A) and calmodulin binding peptide (CBP) epitopes of the original TAP-tag are in light blue, the TEV protease site is black, and spacer sequences are grey. In the PTP-tag, the Protein C (Prot C) epitope is depicted in dark blue. Reprinted with permission from Elsevier: Schimanski *et al.* (2005b), *Eukaryotic Cell* 4, pp 1944.

In consecutive steps, PTP-tagged protein purification is carried out as follows, see Figure 5-42. Firstly, the tagged protein is bound to IgG beads or matrix via the Protein A epitope and multiple washes ensure that unbound proteins (C: contaminants) are removed, and the protein of interest and possible interactors remain. Next the protein or protein complexes are then released from the IgG and Protein A by cleavage with the TEV protease. These released proteins are then bound to anti-Protein C beads and washed again. The proteins were released from the beads using an EDTA buffer, and were eluted three times (Materials and Methods, Section 2.12 for the protocol). The eluate was then run on a 10 % SDS-PAGE gel and probed with anti-PAP anti-serum which detects the ProtA and anti-ProtC anti-serum which detects the ProtC.



**Figure 5-42: Overview of the PTP-purification strategy.**

In consecutive steps, PTP-tagged protein purification was performed as follows: In the first affinity column, the tagged protein is bound to IgG beads via the Protein A (ProtA) epitope. Washes ensure that contaminants (C) are removed. The protein(s) are then released from the IgG bead and Protein A by cleavage with the TEV protease (indicated by the scissors). In the second affinity column, the released protein(s) are then bound to anti-Protein C beads and washed again. The purified protein(s) are released from the beads using an EDTA buffer. Reprinted with permission from Elsevier: Piug *et al.* (2007), *Methods* **24**, pp 219.

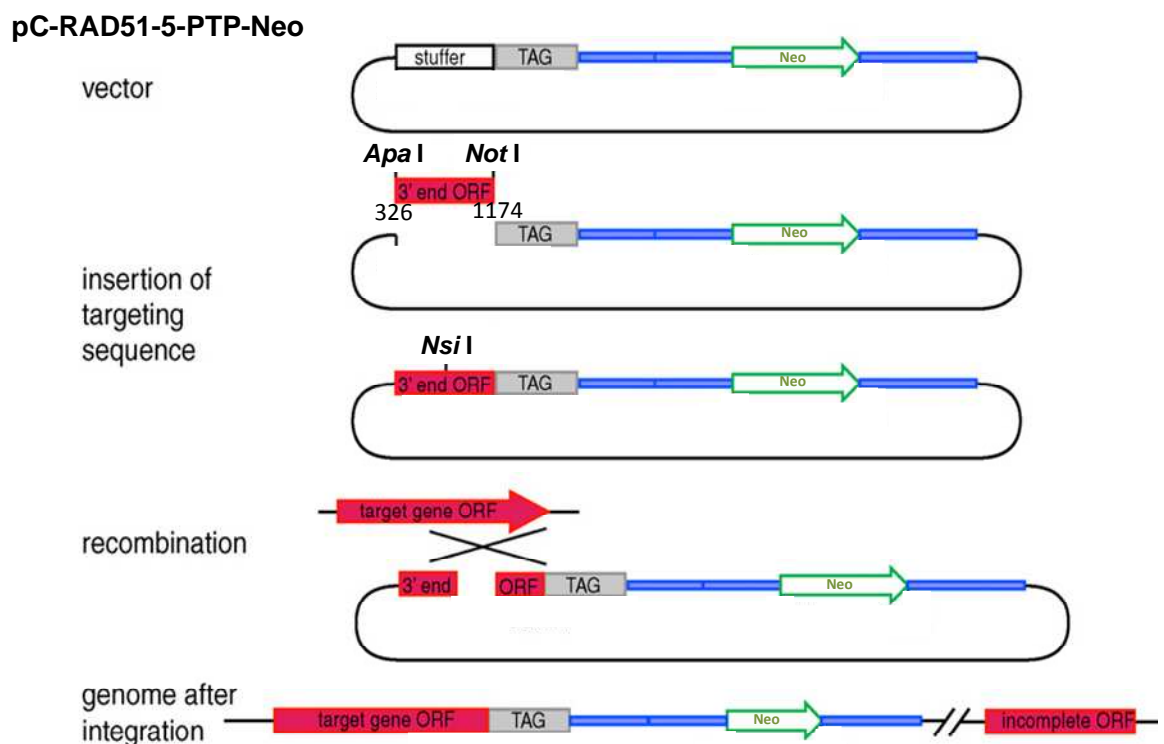
## 5.7.2 Generation of the PTP-tagged RAD51 paralogue constructs

### 5.7.2.1 Generation of the PTP tagged RAD51-5 construct

To generate a C-terminal PTP-tagged RAD51-5 construct, 848 bp of the 3' end of the *RAD51-5* ORF was cloned into the pC-PTP-vector (Schimanski *et al.*, 2005a). This fragment of *RAD51-5* was PCR-amplified using the 51-5 Cterm PTP forward and reverse



primers (see Materials and Methods, Section 2.4.1, Table 2-1), which had the enzyme restriction sites *Apa*I and *Not*I on the forward and reverse primers respectively, to facilitate the cloning into the vector (see Figure 5-43). Once clones were confirmed and sequenced, the construct was linearised using a unique restriction enzyme site in the middle of the cloned fragment of *RAD51-5*, *Nsi*I. This was then transformed into wildtype bloodstream stage Lister 427 *T. brucei*. Recombination of the construct should occur into the endogenous *RAD51-5* locus. This resulted in the whole ORF becoming C terminally fused with the PTP ORF, as well as insertion of the resistance cassette for selection, in this case neomycin (Neo; see Figure 5-43).



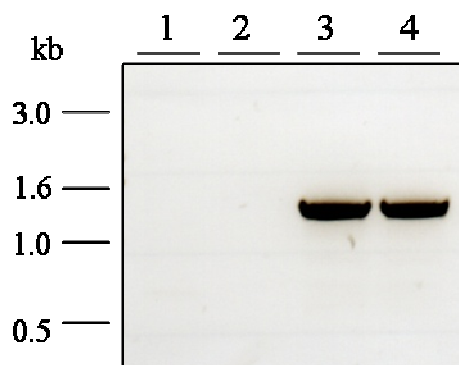
**Figure 5-43: Diagram showing design and use of the C-terminal PTP vector for *RAD51-5* tagging at its endogenous locus**

A fragment from 326 bp to 1174 bp of the *RAD51-5* ORF was cloned into the pC-PTP-vector using *Apa*I and *Not*I restriction enzyme sites. This was then linearised using a *Nsi*I site and transformed into wildtype *T. brucei* cells. Green arrow: selectable resistance cassette, (Neo; G418 resistance); Red arrow: targeted open reading frame, *RAD51-5*. Diagram adapted from Kelly *et al.*, 2007.

### 5.7.2.2 Confirmation of the PTP tagged *RAD51-5* in *T. brucei*

Once *T. brucei* *RAD51-5* PTP transformants had been selected by G418 resistance at a concentration of  $2.5 \mu\text{L} \cdot \text{mL}^{-1}$ , the genomic DNA was extracted and a PCR-amplification check was carried out to confirm that the construct had integrated into the *RAD51-5* endogenous locus. This was carried out using a forward primer which anneals with the 5' end of the *RAD51-5* ORF and a reverse primer which anneals in the middle of the PTP tag: named *RAD51-5* For and PTP-Cterm Check Rev, respectively (see Materials and Methods,

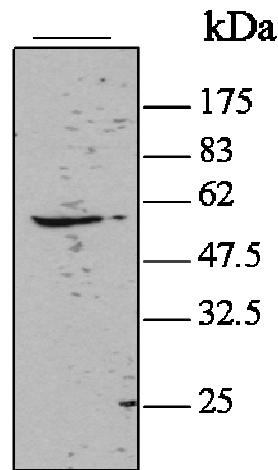
Section 2.4.1, Table 2-1). The total length of PTP tag is 506 bp and the PTP-Cterm Check reverse primer anneals at position 222 bp from the 5' end. Figure 5-42 shows the results of the PCR-amplification from two negative controls [genomic DNA extracted from wildtype untransformed cells (Figure 5-42, Lane 1), and pC-RAD51-5-PTP-Neo plasmid DNA (Figure 5-44, Lane 2)], and two transformed G418-resistant clones (Figure 5-44, Lane 3 and 4). This figure confirmed that the PTP tag had integrated into the endogenous *RAD51-5* locus: the *RAD51-5* ORF is 1174 bp in length, and the amount of the PTP tag ORF PCR-amplified is 222 bp, so the total fragment size is 1396 bp, which is consistent with the size of the product observed.



**Figure 5-44: PCR-amplification to check PTP-tag integration into the *RAD51-5* allele.**

PCR-amplification was carried out using a forward primer which bound to the 5' end of the *RAD51-5* ORF and a reverse primer which bound to the middle of the PTP tag, named *RAD51-5* For and PTP-Cterm Check Rev primers respectively, to confirm that the PTP tag had integrated into the endogenous locus of *RAD51-5*. DNA extracted from two G418 resistant positive transformants was used as template for the PCR reaction (Lane 3 and 4). Genomic DNA extracted from wildtype untransformed cells (Lane 1), and plasmid DNA from the original construct: pC-RAD51-5-PTP-Neo (Lane 2) was used as negative controls.

One of the clones analysed in Figure 5-44 (Lane 3) was chosen for further analysis. The cells were grown to a density of approximately  $3 \times 10^6$  cells per ml, in 25 ml of HMI-9, lysed and the extracted proteins were separated on a 10 % SDS-PAGE gel. These proteins were then transferred to a nylon membrane and probed with anti-PAP antibody conjugated to HRP (Roche) at a dilution of 1:1000. The PTP-tag fused to RAD51-5 adds 19 kDa resulting in a total RAD51-PTP size of 61 kDa. Figure 5-45 shows the whole cell lysate probed with the anti-PAP antibody, revealing a band of the expected size. This confirms the PCR-amplification analysis and suggests that RAD51-5 has been successfully tagged with PTP.

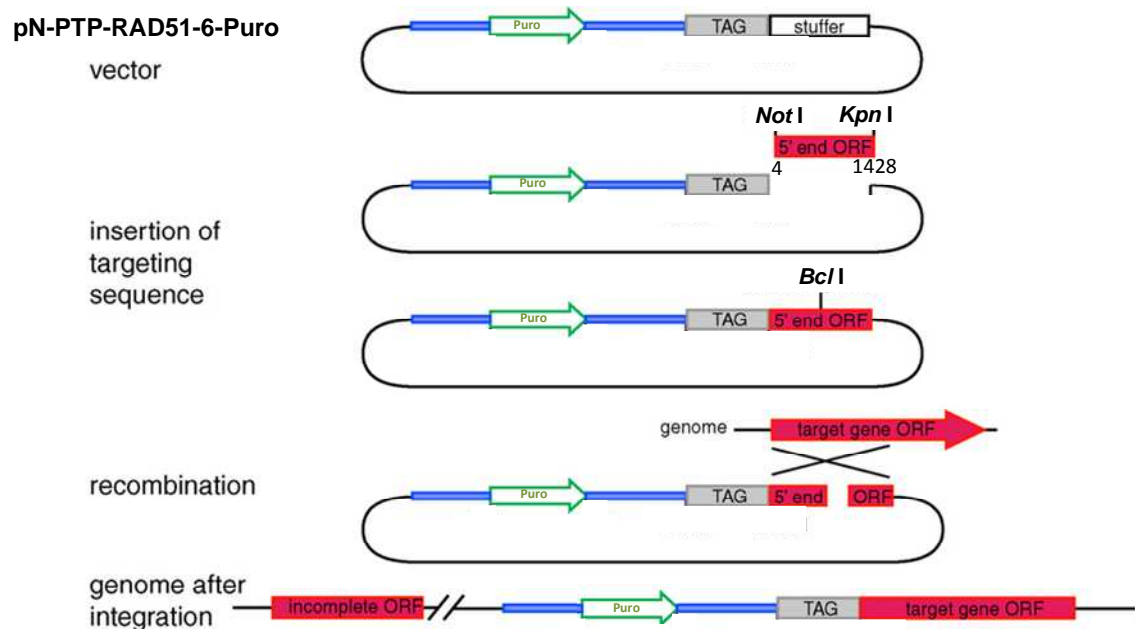


**Figure 5-45: Western analysis to confirm the integration of the PTP-tag into the endogenous RAD51-5 locus.**

The transformed PTP-RAD51-5 cells were lysed and the extracted proteins separated on a 10 % SDS-PAGE gel. These proteins were then transferred to a nylon membrane and probed with anti-PAP antibody conjugated to HRP. The PTP tag fused to the RAD51-5 is 61 kDa in size, and can be seen below the 62 kDa marker band.

### **5.7.2.3 Generation of the PTP-tagged RAD51-6 construct**

To overcome the lack of specificity of the anti-RAD51-6 anti-sera which could not be used for co-immunoprecipitation analysis (see Section 5.5.5), PTP-tagging of *RAD51-6* was attempted. To generate an N-terminal PTP-tagged RAD51-6, the whole ORF of *RAD51-6* was cloned into the pN-PTP-vector. *RAD51-6* was PCR-amplified using 51-6 Nterm PTP forward and reverse primers (see Materials and Methods, Section 2.4.1, Table 2-1). These primers had *NotI* and *KpnI* enzyme restriction sites, on the forward and reverse respectively, and were used to clone the fragment into the vector, see Figure 5-46. The fragment of *RAD51-6* ORF cloned into the vector pN-PTP was from 4 bp to 1428 bp. This pN-PTP-RAD51-6-Puro construct was confirmed by sequencing and linearised using a unique restriction enzyme site in the middle of *RAD51-6*, *BclI*. The linearised construct was transformed into the heterozygous *RAD51-6* mutant, which had one *RAD51-6* allele disrupted with a blasticidin resistance cassette (see Chapter 3, Section 3.5.3). Recombination of the construct into the intact allele of *RAD51-6* should result in the intact ORF becoming N-terminally tagged with PTP, with insertion via selection for the resistance cassette for puromycin, see Figure 5-46.

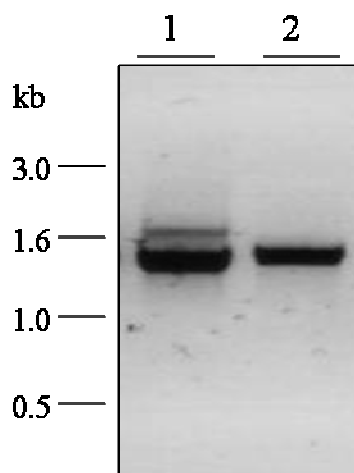


**Figure 5-46: Diagram showing design and use of an N-terminal PTP vector for *RAD51-6* tagging at its endogenous locus.**

A fragment of *RAD51-6* (from 4 bp to 1428 bp of the ORF) was cloned into the pN-PTP vector using *NotI* and *KpnI* restriction enzyme sites. This was then linearised using a *BclI* restriction site and transformed into *T. brucei* *RAD51-6* +/- cells. Green arrow: selectable resistance cassette, puromycin (Puro); Red arrow: targeted open reading frame, *RAD51-6*. Diagram adapted from Kelly *et al.* (2007).

#### 5.7.2.4 Confirmation of PTP tagged *RAD51-6* in *T. brucei*

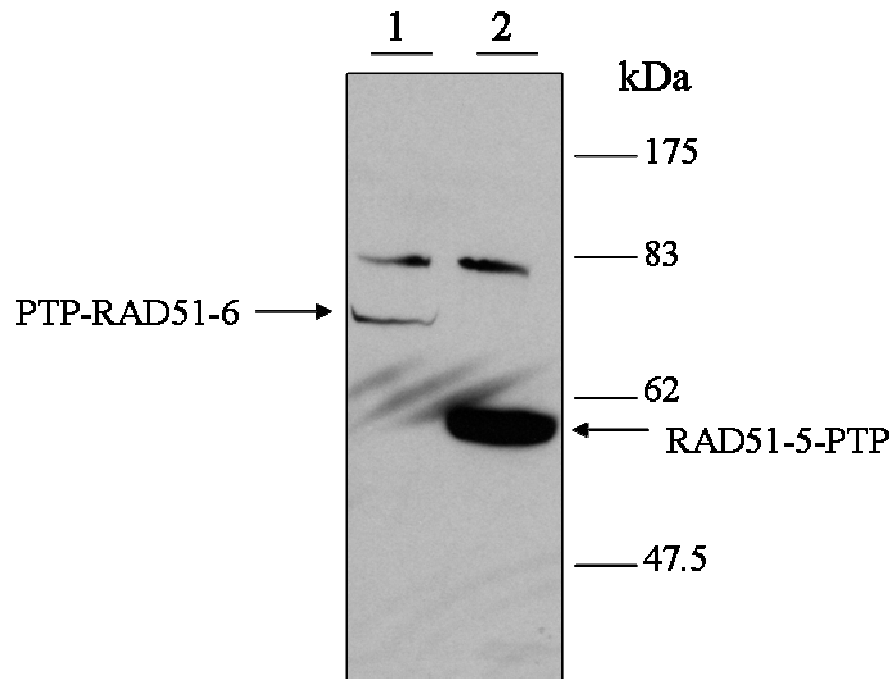
A selection of PTP-*RAD51-6* clones that were resistant to puromycin were analysed to confirm integration of the PTP-tag. Genomic DNA was extracted from the clones and PCR-amplification was carried out. As a positive control, the whole ORF of *RAD51-6* was PCR-amplified using the RECA2 StartORF-For and RECA2 EndORF-Rev (see Materials and Methods, Section 2.4.1, Table 2-1). Figure 5-47 shows the PCR products of one PTP-*RAD51-6* clone. Two fragments were obtained because the PTP-*RAD51-6* construct was transformed into the *RAD51-6* +/- cell line, which had one *RAD51-6* allele disrupted with blasticidin resistance cassette (Chapter 3, Section 3.5.3). This resulted in a larger product as the resistance cassette added approximately 250 bps to the disrupted region (Figure 5-47, Lane 1: 1685 bp fragment). The smaller fragment is the size of the intact *RAD51-6* allele (Figure 5-47, Lane 1: 1435 bp fragment). Using the same extracted genomic DNA and the PTP-Nterm check For and RECA2 EndORF-Rev primers, PCR-amplification was carried out to check if the PTP tag integrated into the *RAD51-6* endogenous locus. Figure 5-47, Lane 2 shows the fragment obtained, confirming that the *RAD51-6*-PTP-tag was present.



**Figure 5-47: PCR-amplification to check PTP-tag integration into the *RAD51-6* allele.**

The genomic DNA was extracted from one puromycin resistant clone and PCR-amplification analysis was carried out to confirm that the PTP-tag had integrated into the *RAD51-6* allele. Lane 1 shows the PCR-products obtained when using RECA2 StartORF-For and RECA2 EndORF-Rev primers. Lane 2 shows the PCR product when PCR-amplification was carried out using a forward primer which bound to the 5' end of the PTP tag and a reverse primer which bound to the 3' end of the ORF of *RAD51-6*.

The protein extracts from this N-PTP-*RAD51-6* clone and the C-*RAD51-5*-PTP clone (as a control) were separated on a 10 % SDS-PAGE gel and then transferred to a nylon membrane (Figure 5-48). The membrane was probed with anti-PAP antibody conjugated to HRP (Roche) at a dilution of 1:1000. The PTP tag fused to the N terminus of *RAD51-6* adds 19 kDa, resulting in a total PTP-*RAD51-6* size of approximately 70 kDa. Figure 5-48 shows this western blot where a band of the predicted size of PTP-*RAD51-6* can be detected. Although there is non-specific binding at approximately 83 kDa in both the protein extracts from the N-PTP-*RAD51-6* and C-*RAD51-5*-PTP clone, a band of the correct size of C-*RAD51-5*-PTP (~61 kDa), although faint, can be detected. This western blot also confirms the PCR-amplification result and shows that the PTP-*RAD51-6* predicted size of 69.4 kDa was present in protein extracts from the transformed clone.



**Figure 5-48: Western analysis showing the PTP-tagging of RAD51-6 by probing the cell lysate with anti-PAP antibody.**

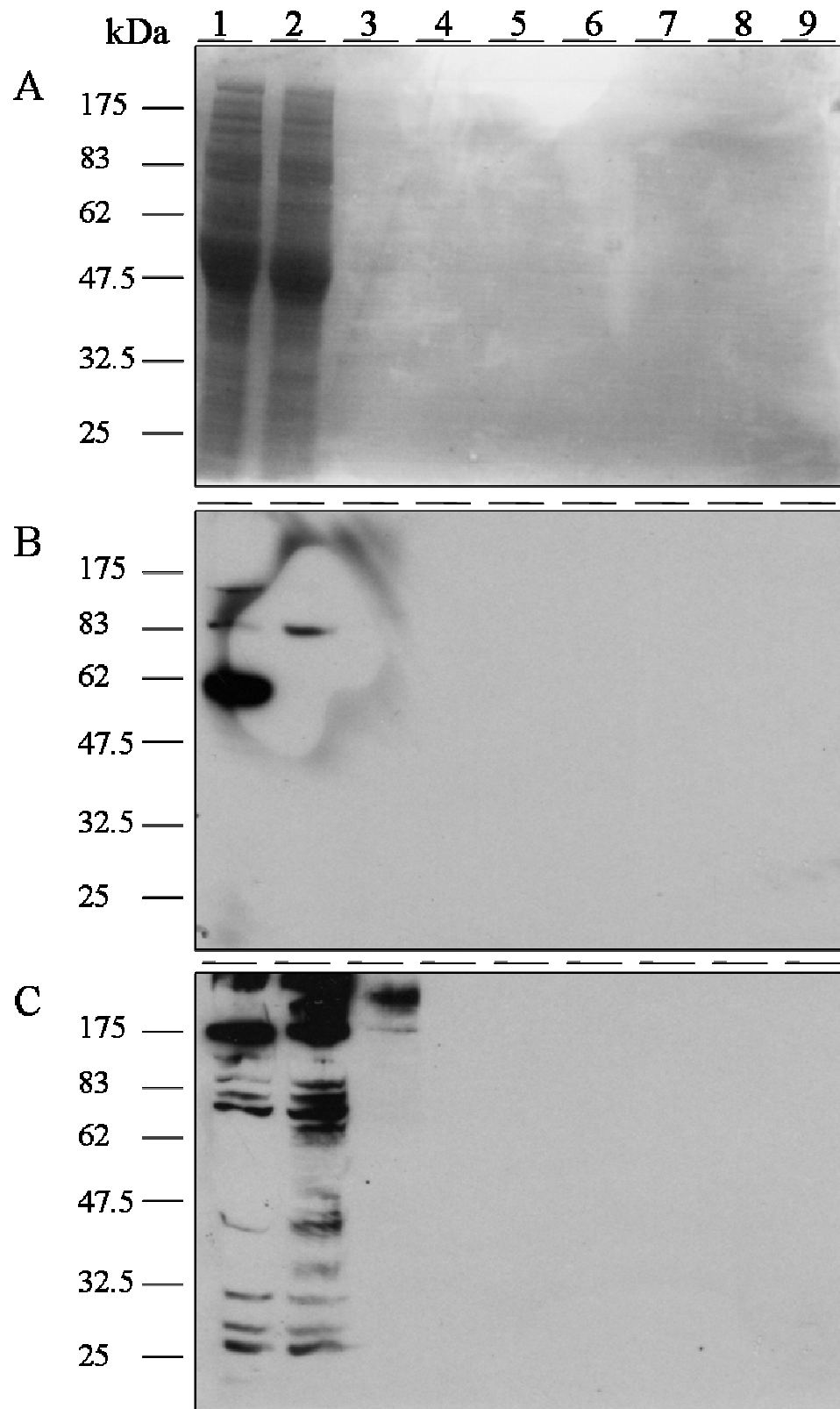
Western blot analysis was carried out to confirm that the PTP-tag fusion of RAD51-6 was correct. Lane 1 has the PTP-RAD51-6 cell extract and Lane 2 has the RAD51-5-PTP cell extract. Both extracts were separated on a 10 % SDS-PAGE gel and transferred to a nylon membrane. This was then probed with anti-PAP antibody conjugated with HRP (at a dilution of 1:1000). The bands at the predicted sizes of PTP-RAD51-6 of ~ 70 kDa (Lane 1) and PTP-RAD51-5 of ~ 61 kDa (Lane 2) are indicated.

### 5.7.3 Attempt to recover protein complexes by PTP purification

Once PTP-tagging was confirmed, an attempt to purify RAD51-5-PTP was performed. Five litres of  $3 \times 10^6$  cells per ml of RAD51-5-PTP cell culture were pelleted and the proteins were extracted as before for the co-IP analysis (see Section 5.6 and Materials and Methods, Section 2.6.2.2). The PTP-tagging purification of RAD51-5-PTP was then carried out, see Figure 5-40 which shows an overview of each step of the purification strategy. The IgG beads were equilibrated in the correct buffers and protease inhibitors were added (see Materials and Methods, Section 2.12 for details). A total of 4 ml of protein extract from the RAD51-5-PTP cells was added to the IgG beads (Invitrogen) and incubated at 4 °C for two hours. This allowed the RAD51-5-PTP protein (target protein) to bind to beads via the Protein A epitope (ProtA). After the incubation, the flow-through was collected and the column was washed, to ensure that all contaminants were removed. The RAD51-5-PTP and possible associated protein(s) remain bound to the IgG matrix. The RAD51-5-PTP was then released from the IgG beads and Protein A by TEV protease cleavage. To do this, 200 units of TEV protease were added to the beads and incubated overnight at 4 °C. After 16 hours, the flow-through containing RAD51-5-ProtC was

collected and then bound to anti-Protein C affinity matrix (Roche). This second affinity column was allowed to incubate at 4 °C for two hours, and then allowed to settle. The flow-through was collected. The matrix was washed before the proteins were released from the beads using an EDTA buffer. The RAD51-5-ProtC and any associated proteins were eluted.

The flow-throughs, washes and elutant from the RAD51-5-PTP purification were electrophoresed on two 10 % SDS-PAGE gels (Figure 5-49, Lanes 1-9). On one gel, the proteins were stained with Coomassie blue (Figure 5-49A), and the other gel, the proteins were transferred to a nylon membrane. The Coomassie stained gel (Figure 5-49A) shows that the majority of proteins were in flow-through from the first IgG affinity column and that the following washes did not contain detectable amounts of protein. The western blot was first probed with anti-PAP-HRP anti-serum (at a dilution of 1:750) which detects the ProtA (Figure 5-49B). This blot showed that the original protein extract contained a band of the predicted size of RAD51-5-PTP (~ 61 kDa) (Figure 5-49B, Lane 1), while in the first flow-through this band was not detected (Figure 5-49B, Lane 2). The RAD51-5-PTP band was also absent in the following washes (Figure 5-49B, Lane 3-4). No other bands were detected in Lanes 5-9; this is expected as the TEV-protease cleavage results in the ProtA remaining bound to the IgG matrix in the column, while releasing the RAD51-5-ProtC and any associated proteins (Figure 5-49B, Lane 5-9). This membrane was stripped and re-probed with anti-ProtC-peroxidase anti-serum (at a dilution of 1:3000; Roche) which detects the ProtC (Figure 5-49C). Unexpectedly, there was no detectable band of the correct size in the RAD51-5-PTP protein extract (~61 kDa), while there was considerable cross-contaminant bands in the extract and in the first flow-through (Figure 5-49C, Lane 1-2). There was also no detectable RAD51-5-ProtC in the final eluate (Figure 5-49C, Lane 9). The western blot shown in Figure 5-47C indicates that the anti-ProtC antibody is not specific and can not detect the RAD51-5-PTP protein. Due to time constraints, it was not possible to attempt to re-purify RAD51-5-PTP and PTP-RAD51-6.



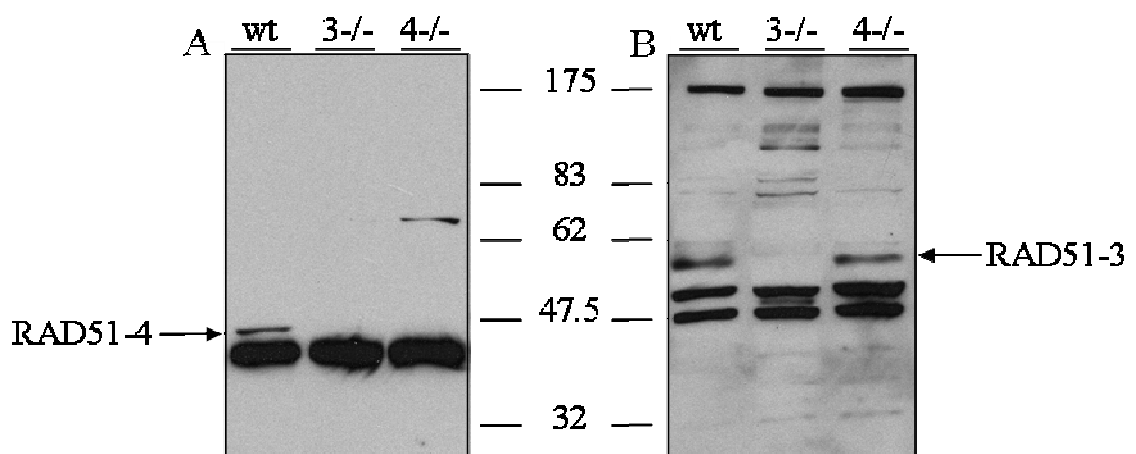
**Figure 5-49: Coomassie stain and western blot of the RAD51-5-PTP purification.**

Two 10 % SDS-PAGE gels were electrophoresed; each had the same solution loaded into the corresponding lanes as follows: Lane 1: RAD51-5-PTP protein extract. Lane 2: Flow-through from the first IgG affinity column. Lane 3: First wash of the IgG affinity column. Lane 4: Final wash of the IgG affinity column. Lane 5: Flow-through after the TEV-protease cleavage. Lane 6: First wash of the anti-ProtC affinity column. Lane 7: Third wash of the anti-ProtC affinity column. Lane 8: Final wash of the anti-ProtC affinity column. Lane 9: The eluate. **Figure 5-49A:** shows the Coomassie stain of one of the gels. **Figure 5-49B:** shows the western blot of the transferred proteins probed with anti-PAP-HRP (at a dilution of 1:750). **Figure 5-49C:** The membrane was stripped and reprobed with anti-ProtC-peroxidase (at a dilution of 1:3000).



## 5.8 Possible co-expression of RAD51-3 and RAD51-4

During the course of testing the specificity of the unpurified and affinity purified anti-sera, it was found that RAD51-4 was absent in the *rad51-3*  $-/-$  mutant protein extracts. In contrast, RAD51-3 was present in the *rad51-4*  $-/-$  mutant protein extracts. This experiment was repeated three times, and the results are shown in Figure 5-50. The wildtype, *rad51-3*  $-/-$  and *rad51-4*  $-/-$  cell extracts were quantified using the Bradford method (Bio-Rad), and  $50 \mu\text{g}.\text{ml}^{-1}$  of each protein extract was separated on a 10 % SDS-PAGE gel. These proteins were transferred to a nylon membrane and probed with affinity purified anti-RAD51-4 anti-serum at a dilution of 1:1000 (see Section 5.5.3). The RAD51-4 was detected with anti-rabbit anti-serum at a dilution of 1:5000 (Molecular Probes). The result of this analysis is shown in Figure 5-50A where the wildtype extract contains a distinct band of the predicted size of RAD51-4 (45.7 kDa) which was not detectable in the two homozygous mutant cell lines, *rad51-3*  $-/-$  and *rad51-4*  $-/-$ . This membrane was stripped by incubation with stripping buffer (Pierce) for 30 minutes, and washed twice in PBS-Tween20 for 30 minutes. The blot was reprobed with affinity purified anti-RAD51-3 anti-sera at a dilution of 1:10 (see Section 5.5.2) and was detected using anti-sheep anti-serum at a dilution of 1:5000 (Santa Cruz). Figure 5-50B shows that RAD51-3 is present in both the wildtype and *rad51-4*  $-/-$  cell extracts, since there was a detectable band of the predicted size of the protein (54.5 kDa). As expected the RAD51-3 band was absent in the *rad51-3*  $-/-$  protein extract (Figure 5-50B).



**Figure 5-50: Western analysis showing possible co-expression of RAD51-3 and RAD51-4.**

Western blot analysis was carried out to confirm the co-expression of RAD51-3 and RAD51-4.  $50 \mu\text{g}.\text{ml}^{-1}$  of protein extract from wildtype, *rad51-3*  $-/-$  and *rad51-4*  $-/-$  cells were separated on a 10 % SDS-PAGE gel and transferred to a nylon membrane. **Figure 5-50A:** shows the membrane when probed with anti-RAD51-4 anti-serum. This was stripped and reprobed. **Figure 5-50B:** shows the blot when probed with anti-RAD51-3 anti-serum. The size markers, the predicted size of RAD51-4 (45.7 kDa) and RAD51-3 (54.5 kDa) are indicated.

There are two possible explanations for this result. First, is that RAD51-4 was not expressed in the *rad51-3*  $-/-$  mutant cell line. The second is that RAD51-3 is required for the stabilisation of RAD51-4 in the cell. RT-PCR-amplification and Northern analysis can be used to determine whether RAD51-4 is transcribed in the *rad51-3*  $-/-$  mutant cell line. This will not conclusively prove that RAD51-4 is translated and therefore not expressed. Another experiment to determine co-expression would be to analyse the *RAD51-3*  $-/+$  re-expressor cell line for the reappearance of RAD51-4.

## 5.9 Summary

The results described in this chapter aim to identify the interactions between the *T. brucei* RAD51 paralogues. It has been shown that the mammalian Rad51 paralogues and yeast Rad51 paralogues complex with one another. The five mammalian Rad51 paralogues have been shown to form two distinct complexes, with one Rad51 paralogue, Rad51C in both complexes (Liu *et al.*, 2002b; Masson *et al.*, 2001b). The two yeast Rad51 paralogues have also been shown to interact with one another (Hays *et al.*, 1995; Johnson & Symington, 1995). Members of each mammalian and yeast complex, XRCC3-Rad51C and Rad55 respectively, have been shown to interact with Rad51 (Hays *et al.*, 1995; Liu *et al.*, 2002b). As a consequence of these findings in other systems, it was predicted that one or more proteins from the *T. brucei* RAD51 family would interact with each other. Since the number of *T. brucei* RAD51 paralogues is similar to that of higher eukaryotes, there was also the potential for more than one complex of RAD51 paralogues. Three approaches were taken to examine if, or which, *T. brucei* RAD51 paralogues interact: yeast-two hybrid, co-immunoprecipitation, and epitope tagging of the RAD51 paralogues.

Using both the yeast-two hybrid and co-immunoprecipitation methods, it has been shown that RAD51-3 and RAD51-4 interact with one another. These two methods provide strong evidence of an interaction between RAD51-3 and RAD51-4 *in vivo*. Yeast-two hybrid analysis has also shown that RAD51-3 and RAD51-6 interact. It is noteworthy that this interaction was weaker than the interaction seen between LexA-RAD51-3 and V5-NLS-B42-RAD51-6. However, the reverse was not seen in the complimentary two-hybrid analysis where LexA-RAD51-6 and V5-NLS-B42-RAD51-3 fusions showed no interaction. This may be because the interaction is not as robust in this system. Nevertheless, substantial interaction between RAD51-3 and RAD51-6 was observed in comparison to the negative control. These results may indicate that, like the mammalian model, two distinct complexes are formed in *T. brucei*, with one RAD51 paralogue, RAD51-3, involved in both complexes. It may also be the case one complex is formed

involving all three proteins, RAD51-3-RAD51-4-RAD51-6. Further analysis is required to define whether or not the *T. brucei* RAD51 paralogues function as multiprotein complexes.

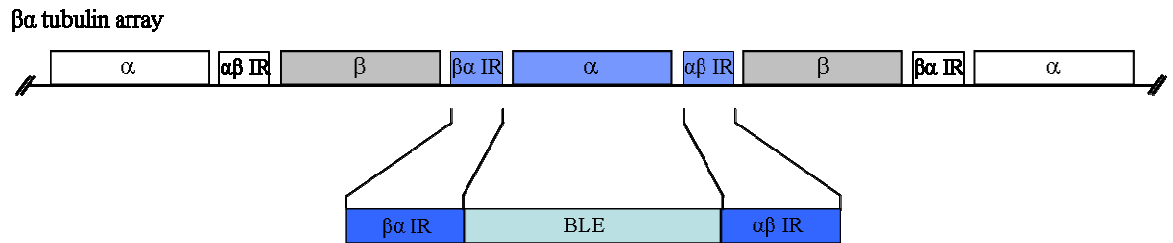
It is widely accepted that Rad51 interacts with itself, forming a nucleoprotein filament on DNA (Benson *et al.*, 1994). Yeast-two hybrid analysis also confirmed that the *T. brucei* RAD51 interacts with itself. In contrast, no other interactions were found between RAD51 and any of the RAD51 paralogues using yeast-two hybrid analysis. It has been shown in Chapters 3 and 4 of this thesis (Sections 3.4.4, 3.7.4 and 4.5) that RAD51 foci formation is decreased in the homozygous mutants of all the RAD51 paralogues: *rad51-3* *-/-*, *rad51-4* *-/-*, *rad51-5* *-/-*, and *rad51-6* *-/-*. Therefore, it is unlikely that RAD51 does not interact with any of the RAD51 paralogues, though it is possible they influence RAD51 function through an unknown intermediate factor. It is conceivable that the RAD51 paralogues do interact with RAD51, even though the analyses used here failed to show interaction. It is possible, for instance, that such interactions are weak, or only occur in the context of damaged DNA.

**Chapter 6. A luciferase assay to  
analyse homologous  
recombination in *T. brucei***

## 6.1 Introduction: Transformation efficiency to analyse homologous recombination rate

To date, analysing the efficiency of homologous recombination (HR) in *T. brucei* has been performed primarily by a transformation efficiency assay by the McCulloch group. This method evaluates the ability of the parasite to undergo HR by measuring the rate of integration of a construct containing an antibiotic resistance marker flanked by sequences that target specific sites of the *T. brucei* genome. The advantage of this approach is that a defined set of targets can be analysed. In most cases, the resistance cassette is flanked with the intergenic processing flanks surrounding the  $\alpha$  tubulin gene, which allows integration into the  $\beta\alpha$ -tubulin array and expression of the drug resistance gene (Conway *et al.*, 2002b; Hartley & McCulloch, 2008; McCulloch & Barry, 1999; Proudfoot & McCulloch, 2005a; Proudfoot & McCulloch, 2005b). This method has also been used to examine the transformation efficiency of different *RAD51* paralogue mutant cells compared with wildtype cells (including in Chapter 3, Section 3.4.3, 3.7.3 and Chapter 4, Section 4.4 of this thesis). These studies demonstrate that integration of the drug resistance gene is under genetic control and consistent with the reaction being driven by HR. It has also been shown that the reaction efficiency is determined by extent of sequence homology between the construct and a single specified target sequence again consistent with it being driven by HR (Barnes & McCulloch, 2007; Bell & McCulloch, 2003).

In this thesis, to measure the transformation efficiency, a linearised piece of DNA containing the ORF for phleomycin resistance (*BLE*), flanked with  $\beta\alpha$  and  $\alpha\beta$  intergenic regions, was transformed into *T. brucei* cells by electroporation (see Materials and Methods, Section 2.8.3). If HR can occur, the targeting regions of the  $\beta\alpha$  and  $\alpha\beta$  intergenic flanks enable the phleomycin resistance cassette to integrate and express from the  $\beta\alpha$ -tubulin array (see Figure 6-1), and as result, the cells become resistant to phleomycin. The tubulin genes are arranged as a tandem array of alternating  $\alpha$  and  $\beta$  genes with a basic repeat length of  $\sim 3.7$  kb. It is thought that this cluster has up to 19  $\alpha\beta$  repeats in *T. brucei* strain 427 (Berriman *et al.*, 2005; geneDB, 2009) and disruption of one  $\alpha$  tubulin gene of the array does not appear to affect the growth rate or morphology of the cells. It has also been shown that the construct is very predominantly integrated by HR (Barnes & McCulloch, 2007).



**Figure 6-1: Transformation efficiency assay**

A linearised piece of DNA containing an antibiotic resistance gene ORF, in this case encoding phleomycin resistance (*BLE*) flanked by  $\beta\alpha$  and  $\alpha\beta$  intergenic regions, is transformed into *T. brucei* cells. If homologous recombination can occur, the flanks allow the resistance cassette to integrate into the  $\beta\alpha$  tubulin array, replacing an alpha ( $\alpha$ ) tubulin ORF (blue box). This allows expression of the phleomycin resistance gene (*BLE*).

Despite the validity of this approach, it measures transformation rather than directly measuring intra-chromosomal or inter-chromosomal recombination, and may be prone to experimental fluctuation. Transformation rate may be affected by many factors including DNA concentration, transformation conditions (for example: fluctuations in delivering voltage during electroporation) and the growth of the cells (for example: mutants in HR also have an increase in population doubling time). It has been shown that *T. brucei* recombination efficiency is dependent on the similarity of the sequence of the targeting regions (Barnes & McCulloch, 2007). In addition, this method uses a relatively non-physiological substrate; the construct is a small, linear, non-replicative, non-transcribed piece of DNA that is not coated with chromatin and must pass through the cell and nuclear membrane to be recombined. The fact that the construct is linear causes this to be an artificial means of testing recombination, as the molecule ends may be viewed as double strand breaks (DSB) which may not always be the lesion that induces recombination. The antibiotic selection may also influence the number of transformants recovered. In separate experiments in this thesis, for instance, transformation efficiency was approximately 4 fold higher for the wildtype cells analysed in Chapter 3 (Section 3.4.3 and Section 3.7.3) compared to the same cell line analysed in Chapter 4 (Section 4.4).

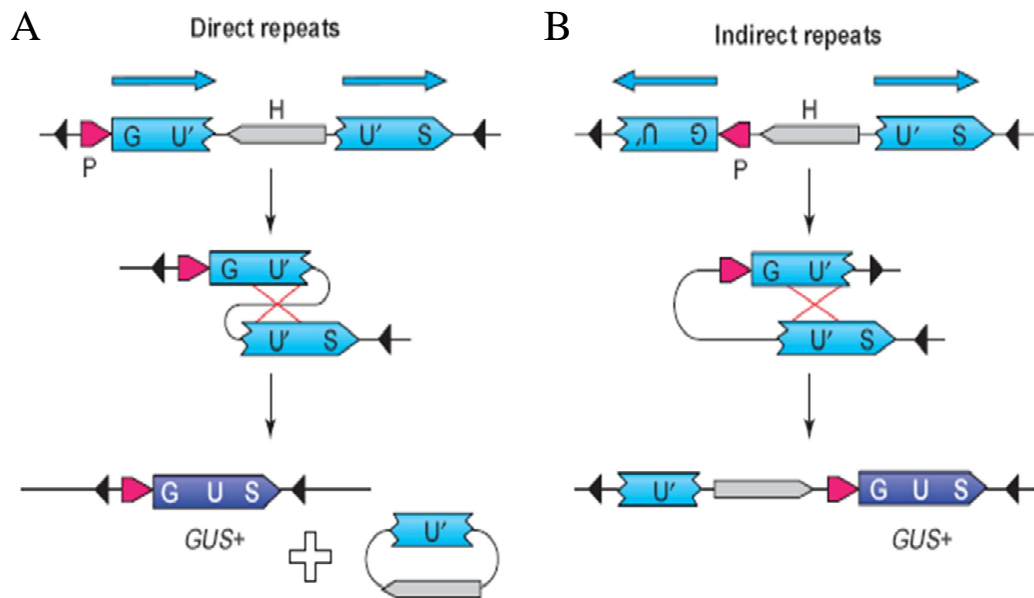
An alternative system has been developed to measure recombination in *T. brucei* by Horn and colleagues. This system uses controlled expression of a nuclease, I-SceI (Glover *et al.*, 2007). I-SceI is a rare cutting endonuclease and, when expressed *in vivo*, a DSB will result at the I-SceI site. After induction of the expression of I-SceI and DSB formation, the cells are allowed to recover, and are measured. The mapping of the survivors provides information regarding the rate and pathway of recombination. Again, however, this cannot analyse the natural rates of HR, or other recombination pathways in *T. brucei*, as it relies on a specific DSB initiation. The objective of this chapter of the thesis was to develop an

alternative assay to test for HR rates and pathways in *T. brucei*, one which is more relevant physiologically and adaptable for different purposes.

## 6.2 Principles of marker constructs to measure homologous recombination

One method of analysing the rate of HR in a non-selective assay system was described by Swoboda *et al.* (1994). The system enabled visualisation of intrachromosomal HR events throughout the whole life-cycle of *Arabidopsis thaliana* by histochemical staining of the whole plant. The system used a disrupted  $\beta$ -glucuronidase (*uidA*) gene as a genomic recombination substrate. Two  $\beta$ -glucuronidase gene fragments sharing a region of sequence homology were integrated into the *A. thaliana* genome. The homologous regions were positioned in the same orientation to one another as direct repeats (see Figure 6-2A: Schuermann *et al.*, 2005) and were separated by a hygromycin phosphotransferase resistance marker. If a recombination event occurred, the two gene fragments would form a functional  $\beta$ -glucuronidase ORF by recombining the homologous regions of the gene fragments (Swoboda *et al.*, 1994). This assay system, called the GUS assay, allowed quantitative and precise localisation of the recombination event by showing functional  $\beta$ -glucuronidase with addition of its substrate, 5-bromo-4-chloro-3-indoly glucuronide (X-Glu) (Swoboda *et al.*, 1994).

This system was modified to detect any HR events in living tissues of *A. thaliana* (Fritsch *et al.*, 2004) by using the firefly luciferase gene (LUC) as the reporter instead of the  $\beta$ -glucuronidase gene. This system contained two fragments of the luciferase gene and, like the GUS assay, each fragment had homologous overlapping regions, see Figure 6-2B. Although Figure 6-2B uses the GUS assay to represent the system, the principle was the same for the LUC assay. The two homologous fragments were positioned as inverted repeats relative to one another. If HR occurred, the result would be a functional luciferase gene. To detect recombinants, the plants were sprayed with luciferin, and bioluminescence was observed (Fritsch *et al.*, 2004).



**Figure 6-2: Intra-chromosomal homologous recombination reporter constructs with direct and indirect repeats.**

Both constructs consist of two inactive fragments of a reporter gene, the  $\beta$ -glucuronidase gene. The gene fragments are GU' and U'S, which share the homologous sequence, U'. During homologous recombination a strand exchange event can occur, which results in the production of the functional reporter gene, GUS. **Figure 6-2A:** shows the recombination event (indicated by the red X) required when the homologous regions are orientated as direct repeats to one another, resulting in the deletion of the sequence between the two fragments. **Figure 6-2B:** shows the recombination event required when the homologous regions are orientated as indirect repeats, resulting in the conservation of the intervening sequence. Reprinted with permission from Elsevier: Schuermann *et al.* (2005), Trends in Genetics **21**, pp 176.

In mammalian cells, similar constructs have been used to assay for HR. Pierce *et al.* (1999) used two modified non-functional GFP genes: one had an I-SceI endonuclease site and two in-frame stop codons within the ORF and the other GFP was truncated at the 5' end. The GFP genes were oriented as direct repeats and were separated with a puromycin N-acetyltransferase gene. I-SceI when expressed *in vivo*, caused a DSB at the I-SceI site (Pierce *et al.*, 1999). The authors showed by flow cytometry that expression of I-SceI in wildtype transfected cell lines produced fluorescence (Pierce *et al.*, 1999). Therefore, the DSB was repaired by HR between the two non-functional GFP genes, resulting in a functional GFP gene. It was also shown in this study that a Rad51 paralogue mutant, *xrcc3*  $-/-$ , did not produced fluorescence compared with the wildtype cells and the authors concluded that *xrcc3*  $-/-$  mutant cells were defective in HR (Pierce *et al.*, 1999).

In other experiments using mammalian cells, disrupted resistance cassettes have been used instead of bioluminescent or fluorescent markers. The puromycin N-acetyltransferase and neomycin phosphotransferase genes have been used by Johnson *et al.* (1999) and Lio *et al.* (2004) respectively. Both studies used two non-functional resistance genes; one was disrupted by addition of an I-SceI site and the other was truncated. Expressing the I-SceI

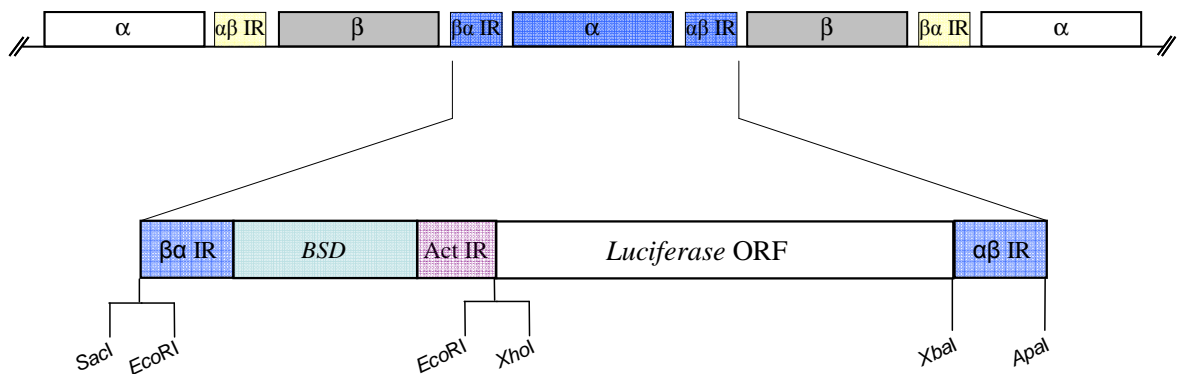


endonuclease in the cells resulted in a DSB in the integrated construct and repair produced a functional resistance gene. The viable cells resistant to the drug were therefore able to undergo HR of the two non-functional drug resistance genes (Johnson *et al.*, 1999; Lio *et al.*, 2004). These studies showed that two of the mammalian Rad51 paralogues, XRCC2 and Rad51C, promote the repair of the DSB by HR (Johnson *et al.*, 1999; Lio *et al.*, 2004).

### 6.3 Generation of a positive control expressing luciferase

Before generating the constructs to assay for recombination, it was necessary to confirm that the firefly luciferase gene expressed in *T. brucei* cells can result in the production of measurable bioluminescence similar to the observations by Wang and colleagues (Sommer *et al.*, 1992). To this end, a positive control construct was generated. This construct was required to allow integration of the luciferase ORF into a transcribed region of the *T. brucei* genome, in particular wildtype bloodstream stage cells Lister strain 427. This construct contained the ORF of luciferase, and blasticidin S deaminase (*BSD*) flanked by the  $\beta\alpha$  tubulin and actin intergenic regions ( $\beta\alpha$  IR and Act IR). To generate this positive control construct, the  $\beta\alpha$  IR, *BSD*, Act IR were cloned into the TOPO vector 2.1, as used in Chapter 3, Section 3.2.2. Briefly, blasticidin (400 bp) and the intergenic processing flanks ( $\beta\alpha$  IR (240 bp) and Act IR (400 bp)) were PCR-amplified from the plasmid pCP101 (Proudfoot & McCulloch, 2005a) using the primers: Tubulin For and Actin Rev (see Materials and Methods, Section 2.4.1, Table 2-1). The resulting PCR product of approximately 1040 bp was cloned into the TOPO 2.1 vector (Invitrogen, see Materials and Methods, Section 2.4.7), resulting in the plasmid TOPO- $\beta\alpha$  IR-*BSD*-Act IR. The luciferase ORF and  $\alpha\beta$  IR were then cloned into the TOPO- $\beta\alpha$  IR-*BSD*-Act IR vector. To this end, the firefly luciferase ORF was obtained from the pGL3-Basic vector (Promega) and, using *Xho*I and *Xba*I sites, the luciferase ORF was digested out of pGL3. The  $\alpha\beta$  IR region was then PCR-amplified from genomic DNA from wildtype *T. brucei* cells, using the AB forward and AB reverse primers, which had *Xba*I and *Apa*I restriction sites (see Materials and Methods, Section 2.4.1, Table 2-1). Using a three way ligation, the luciferase ORF and  $\alpha\beta$  IR were cloned into TOPO- $\beta\alpha$  IR-*BSD*-Act IR using the *Xho*I and *Apa*I restriction sites, with the *Xba*I site in between the luciferase ORF and  $\alpha\beta$  IR fragments. The resulting construct is as follows: TOPO- $\beta\alpha$  IR-*BSD*-Act IR-LUC- $\alpha\beta$  IR. Once confirmed by DNA sequenced, the construct was then linearised using *Sac*I and *Apa*I restriction sites and transformed into wildtype *T. brucei* cells (see Materials and Methods, Section 2.1.2). Integration of the construct should occur using the targeting regions ( $\beta\alpha$  IR and  $\alpha\beta$  IR).

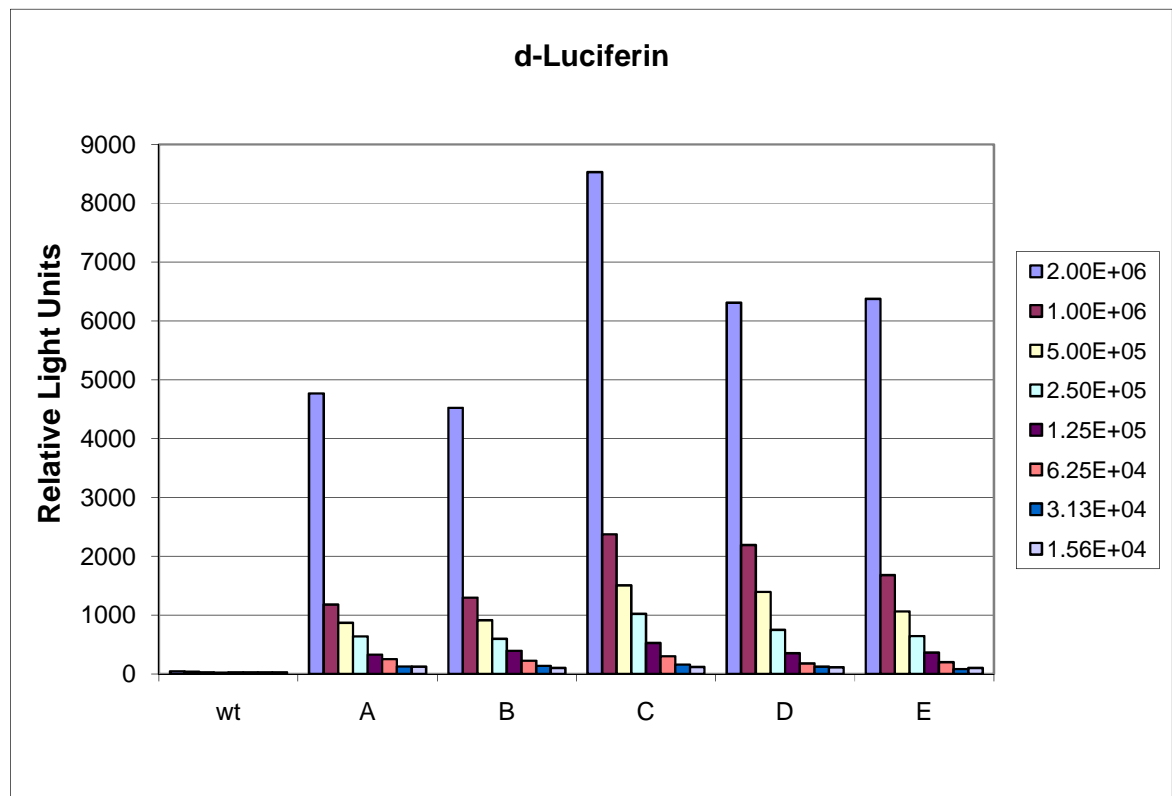
recombining it into the  $\beta\alpha$ -tubulin array with the replacement of one  $\alpha$ -tubulin gene (Figure 6-3).



**Figure 6-3: The positive control construct allowing luciferase expression from the  $\beta\alpha$ -tubulin array.**

The luciferase open reading frame (ORF) and  $\alpha\beta$ -tubulin intergenic region ( $\alpha\beta$ -IR) were cloned with the blasticidin resistance cassette (*BSD*) and intergenic processing flanks,  $\beta\alpha$ -tubulin intergenic region ( $\beta\alpha$  IR) and actin intergenic region (Act IR). This construct was digested and linearised using restriction enzymes *SacI* and *ApaI* and integrated into the  $\beta\alpha$ -tubulin array.

Once electroporated, transformants were selected for by growth in media containing 2.0  $\mu\text{g}\cdot\text{ml}^{-1}$  of blasticidin (Calbiochem; see Materials and Methods, Section 2.1.2). Five clones (A-E) were then selected to assay for luciferase activity (see Materials and Methods, Section 2.13). To do this, the cell density of each was counted, and using a 96 well plate, the five transformants were serially diluted from a density of  $1 \times 10^7$  cells in PBS to final densities of  $2 \times 10^6$  to  $1.56 \times 10^4$  cells in a total volume of 200  $\mu\text{l}$ . To each well, 20  $\mu\text{l}$  of 1 mM of d-luciferin/PBS (Promega) was then added, and incubated at 37 °C for 30 minutes. After this, the plate was removed from the incubator and left at room temperature for 15 minutes. Bioluminescence was then read using a 404 nm filter (Envision). Figure 6-4 compares the bioluminescence emitted from wildtype, non-transformed cells and the five blasticidin resistant transformants.



**Figure 6-4: Bioluminescence production in *T. brucei* bloodstream stage cells.**

The cells from wildtype and five blasticidin resistance transformants (A-E) were serially diluted from  $1 \times 10^7$  cells to  $7.8 \times 10^4$  cells per ml in a final volume of 200  $\mu$ l, to assay for luciferase activity. The total number of cells analysed are indicated in the box on the right. This was carried out with the addition of the luciferase substrate, d-luciferin, and bioluminescence was measured in relative light units.

Bioluminescence generated by luciferase was measured in an enzymatic reaction where d-luciferin is converted to oxyluciferin and light (see review: Viviani, 2002). All five transformants produced bioluminescence in the presence of substrate, whereas no such reaction was detected with the non-transformed wildtype cells. This shows that as few as  $2.5 \times 10^5$  cells expressing luciferase can be detected by light emission (blue points in Figure 6-2), since this was clearly higher than the background readings for the wildtype cells. This suggests that *T. brucei* strain 427 is able to express luciferase from the  $\beta\alpha$ -tubulin array and that such expression can be detected at low densities. This assay also confirmed that d-luciferin is not toxic to the cells as recovery of all five transformants was possible.

## 6.4 A direct repeat luciferase construct to assay homologous recombination

### 6.4.1 Generation of the luciferase construct

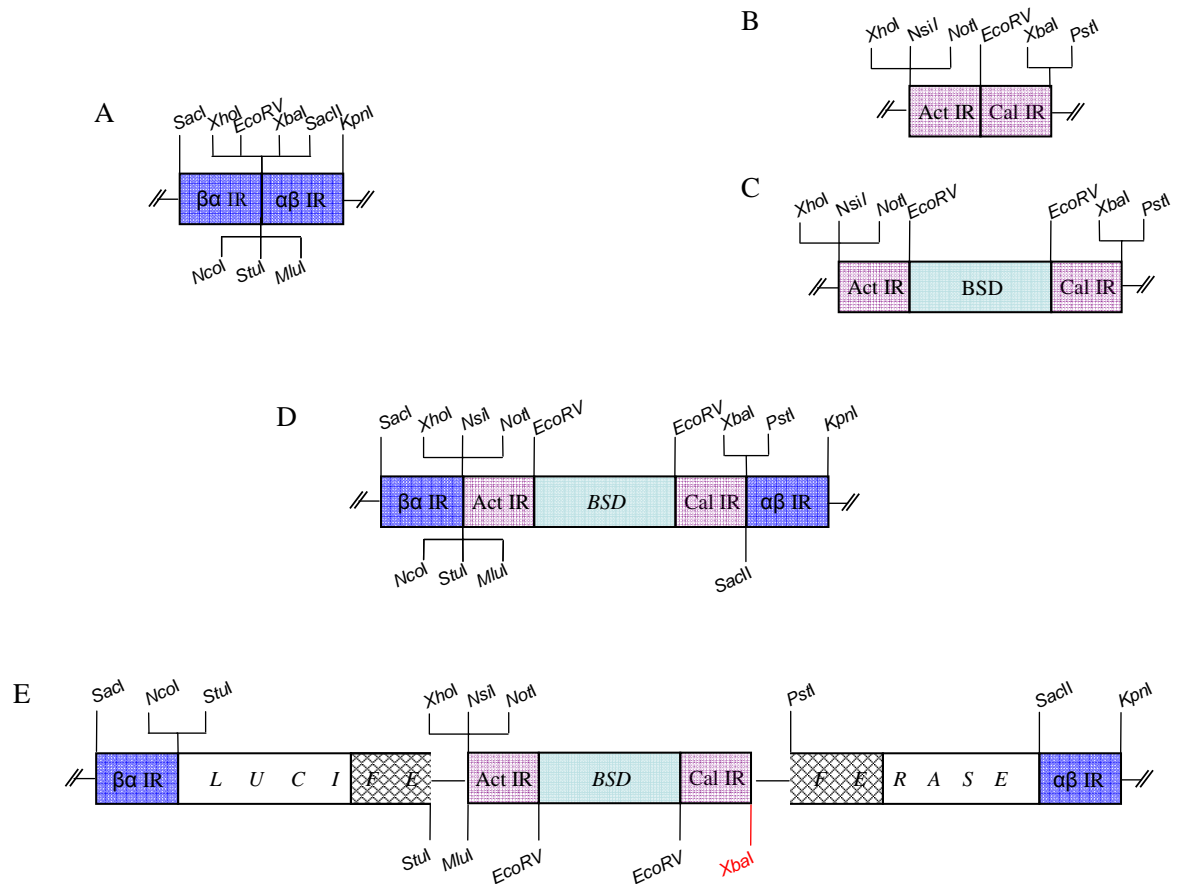
Since the positive control construct, where the luciferase ORF was expressed from the  $\beta\alpha$ -tubulin array, showed that bioluminescence could be detected in bloodstream form cells, a luciferase reporter construct to assay for HR was generated. The first reporter construct contained two inactive fragments of the luciferase reporter gene (see Figure 6-5E). As for similar assays in other cells (see Section 6.2), these fragments contained homologous sequence (shaded area in Figure 6-5E) at their 3' and 5' ends, and were cloned in the same orientation to one another as direct repeats. The *LUC* gene fragments were separated by an antibiotic resistance cassette, phleomycin (*BLE*).

This construct was generated in three parts (Figure 6-5). The first part consisted of cloning the  $\beta\alpha$  and  $\alpha\beta$  tubulin intergenic regions into pBluescript. The fragment of the  $\beta\alpha$  intergenic region was PCR-amplified using primers BA IR For and BA IR Rev, and the  $\alpha\beta$  intergenic region was PCR-amplified using primers AB IR For and AB IR Rev (see Materials and Methods, Section 2.4.1, Table 2-1). These fragments were cloned separately into TOPO vector 2.1 (Invitrogen) and sequenced. Both  $\beta\alpha$  and  $\alpha\beta$  intergenic regions were then digested out of the TOPO vector using *SacI* and *EcoRV*, and *EcoRV* and *KpnI*, respectively, then cloned together into pBluescript. To do this, a three way ligation between these two fragments and *SacI* and *KpnI* digested pBluescript was performed, which resulted in the reconstitution of an *EcoRV* site between the two fragments. The resulting construct was named pBluescript- $\beta\alpha$  IR: *EcoRV*:  $\alpha\beta$  IR (see Figure 6-5A).

The second stage required another pBluescript plasmid into which the actin and calmodulin intergenic regions were cloned. The actin intergenic region was PCR-amplified using the primers Actin For and Rev, and the calmodulin intergenic region was PCR-amplified using the primers Calmodulin For and Rev (see Materials and Methods, Section 2.4.1, Table 2-1). Again these fragments were cloned separately into the TOPO vector 2.1 (Invitrogen) and then digested out of TOPO using *XhoI* and *EcoRV* for the actin intergenic region, and *EcoRV* and *XbaI* for the calmodulin intergenic region. The fragments were then cloned, using a three way ligation, into pBluescript pre-cut with *XhoI* and *XbaI*. As a result, an *EcoRV* restriction enzyme site was generated between the two fragments, and the construct was named pBluescript-Act IR: *EcoRV*: Cal IR, see Figure 6-5B. Next, the phleomycin resistance ORF (*BLE*) was PCR-amplified using the primers Bleo For and Bleo Rev, and

cloned into the pBluescript-Act IR: *EcoRV*: Cal IR construct. The *BLE* ORF was cloned between the actin and calmodulin intergenic processing regions using the *EcoRV* site. The resulting construct pBluescript-Act IR: *BLE*: Cal IR (see Figure 6-5C) was sequenced. Finally, the phleomycin resistance cassette comprising the *BLE* ORF, Act IR and Cal IR processing flanks was digested out of pBluescript-Act IR: *BLE*: Cal IR using the *XhoI* and *XbaI* restriction sites and cloned into *EcoRV*-digested pBluescript- $\beta\alpha$  IR: *EcoRV*:  $\alpha\beta$  IR, with the removal of the *EcoRV* restriction enzyme site between the  $\beta\alpha$ : $\alpha\beta$  intergenic regions. The resulting construct was pBluescript- $\beta\alpha$  IR-Act IR-*BLE*-Cal IR- $\alpha\beta$  IR (see Figure 6-5D).

The third part of the generation of the LUC construct involved the cloning of the two putatively inactive fragments of the luciferase gene, which contained overlapping homologous sequence. The fragments were cloned into pBluescript- $\beta\alpha$  IR-Act IR-*BLE*-Cal IR- $\alpha\beta$  IR. The start of the luciferase ORF was PCR-amplified using the primers Start Luc For and Rev which amplified 1064 bp of the 5' end of the gene ('*LUCIFE*' in Figure 6-5E). The end of the luciferase ORF was PCR-amplified using the primers End Luc For and Rev which amplified 925 bp of the 3' end of the gene ('*FERASE*' in Figure 6-5E). 336 bp is shared in the two fragments ('*FE*' in Figure 6-5E). The 5' end of the luciferase gene was cloned using the *StuI* restriction site of the pBluescript- $\beta\alpha$  IR-Act IR-*BLE*-Cal IR- $\alpha\beta$  IR between the  $\beta\alpha$  IR and Actin IR, and the orientation was confirmed. Then the 3' end of the luciferase gene was cloned using the *PstI* and *SacII* restriction sites with the loss of the *XbaI* site (indicated in red in Figure 6-5E). The resulting construct was pBluescript- $\beta\alpha$  IR-LUCIFE-Act IR-*BLE*-Cal IR-FERASE- $\alpha\beta$  IR (see Figure 6-5E).



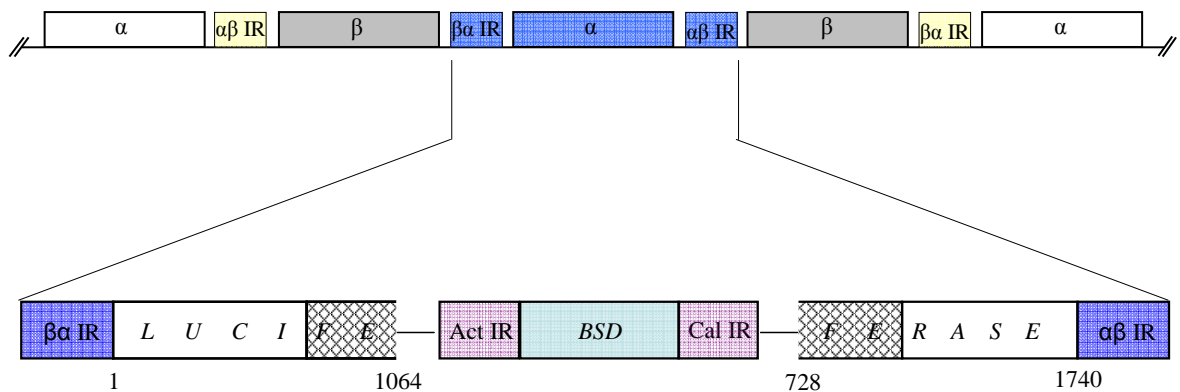
**Figure 6-5: Generation of a direct repeat luciferase construct to assay homologous recombination.**

The diagram shows the cloning strategy for the luciferase construct. **Figure 6-5A:** The  $\beta\alpha$  and  $\alpha\beta$  tubulin intergenic regions ( $\beta\alpha$  IR and  $\alpha\beta$  IR) were cloned into pBluescript. **Figure 6-5B:** The actin IR (Act IR) and calmodulin IR (Cal IR) intergenic regions were cloned into a second Bluescript vector, **Figure 6-5C:** shows the integration of phleomycin resistant cassette (BLE) between the two actin IR and calmodulin IR. **Figure 6-5D:** shows the integration of the phleomycin resistance cassette and processing flanks, actin IR and calmodulin IR, in between the  $\beta\alpha$  IR and  $\alpha\beta$  IR, resulting in pBluescript- $\beta\alpha$  IR-Act IR-BLE-Cal IR- $\alpha\beta$  IR. **Figure 6-5E:** The two luciferase fragments (5' end: LUCIFE; and 3' end: FERASE) have homologous sequence (shaded area; FE), were cloned into pBluescript- $\beta\alpha$  IR-Act IR-BLE-Cal IR- $\alpha\beta$  IR. The restriction sites are shown and *XbaI* in red indicates that this site has been removed during the cloning process.

## 6.4.2 Transformation of the construct into *T. brucei* cells

The plasmid was linearised using *SacI* and *KpnI* restriction enzymes and the DNA was transformed into wildtype *T. brucei* bloodstream stage Lister strain 427 cells. The  $\beta\alpha$  and  $\alpha\beta$  intergenic sequences allowed integration into the  $\beta\alpha$ -tubulin array, replacing the  $\alpha$  tubulin gene (Figure 6-6). The actin and calmodulin intergenic regions provide the processing flanks for expression of the phleomycin resistance gene. Transformants were selected with  $1.0 \mu\text{g}.\text{ml}^{-1}$  of phleomycin. The two luciferase fragments are oriented in direct repeat and share of 336 bp of homologous sequence (shaded area in Figure 6-6). Previous work has shown that this is a sufficient length of sequence for HR to occur (Barnes & McCulloch, 2007). Once transformed, passage of the transformants in medium containing phleomycin ensures the fragments of the luciferase gene will not recombined,

as recombination within the luciferase gene fragments would result in the loss of the phleomycin resistance gene (see Figure 6-2A).



**Figure 6-6: Integration of the luciferase construct into the  $\beta\alpha$  tubulin array.**

The construct should integrate into the  $\beta\alpha$ -tubulin array by homologous recombination on the  $\beta\alpha$  and  $\alpha\beta$  tubulin intergenic region ( $\beta\alpha$  IR /  $\alpha\beta$  IR), replacing an alpha ( $\alpha$ ) tubulin gene. The actin IR (Act IR) and calmodulin IR (Cal IR) processing flanks surround the phleomycin resistant ORF (*BLE*) which allows selection for transformants. The two luciferase fragments (*LUCIFE* and *FERASE*) share 336 bp of homologous sequences (shaded; *FE*).

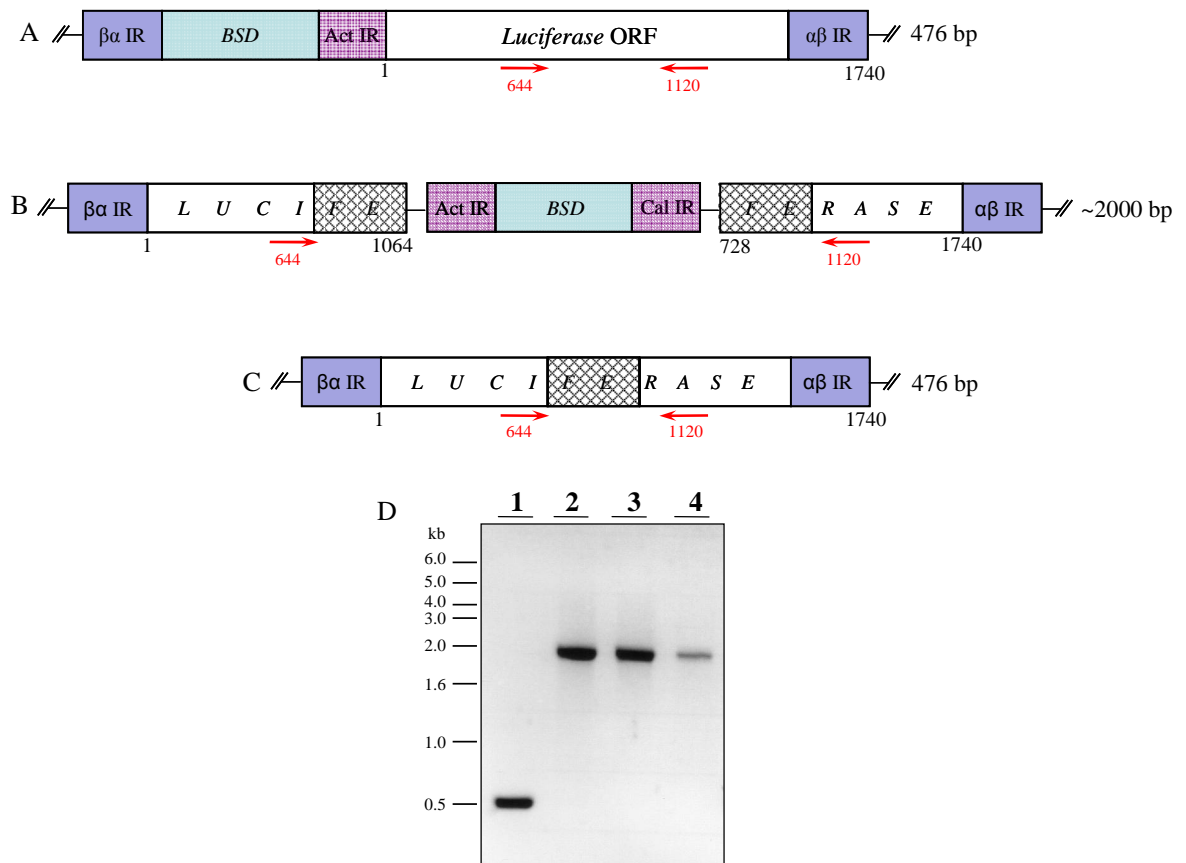
### 6.4.3 Assaying for homologous recombination using the luciferase construct with homologous regions as direct repeats

Two phleomycin resistant transformants (1 and 2) were analysed. For each clone, two passages were performed; one remained on the phleomycin selection and the other without selection. Removal of the phleomycin selection should allow recombination of the homologous ends of luciferase to occur, generating a functional luciferase gene whose gene product could be detected by bioluminescence.

Initial experiments tested whether the potential recombination events could be detected by PCR-amplification with primers that anneal to sequences within the boundaries of the integrated construct: Recombine Check For and Rev (see Materials and Methods, Section 2.4.1, Table 2-1). By analysing the sizes of the PCR-amplified, it could be ascertained whether or not HR had occurred (Figure 6-7). The primer Recombine Check For anneals upstream of the homologous sequence at position 644 bp of the *LUC* ORF and the primer Recombine Rev Check anneals downstream of the homologous sequence at 1120 bp. For the positive control, genomic DNA from clone A (see Section 6.3, Figure 6-4), which had the intact luciferase ORF integrated into the  $\beta\alpha$  tubulin array, was used as a template (Figure 6-7A and Figure 6-7D, Lane 1). The primer annealing sites are 476 bp apart in the intact luciferase ORF. Two clones (1 and 2) transformed with the  $\beta\alpha$  IR-*LUCIFE*-Act IR-BLE-Cal IR-*FERASE*- $\alpha\beta$  IR construct were continually passaged on phleomycin. Genomic

DNA from them was used as template and the resulting PCR products are shown in lanes 2, and 3 (Figure 6-7D). As recombination of the luciferase homologous sequence was prevented by phleomycin retention, the primer binding sites were further apart compared to the positive control (Figure 6-7B), generating a product of approximately 1980 bp. One clone transformed with the construct  $\beta\alpha$  IR-*LUCIFE*-Act IR-BLE-Cal IR-*FERASE*- $\alpha\beta$  IR was taken off phleomycin selection (Clone 1), and passaged for three months without the drug. After this period, genomic DNA was extracted from approximately  $5 \times 10^6$  cells and PCR was performed using the same Recombine Check For and Rev. If recombination of the luciferase fragments occurred in this clone, the PCR product obtained would be the same size as that amplified from the positive control (Figure 6-7C). However, the PCR product obtained was ~ 2.0 kb, the same as the clones passaged on phleomycin (Figure 6-7D, Lane 4), and no PCR product was observed at 476 bp. Therefore no recombination was detected.

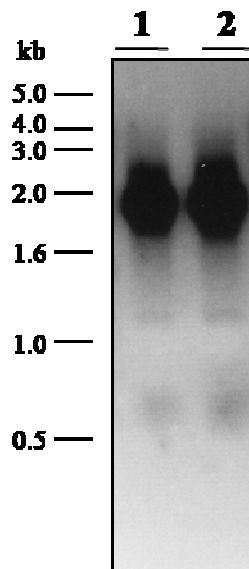




**Figure 6-7: Analysing recombination of the luciferase homologous sequence by PCR-amplification.**

The Recombine Check For and Rev approximate primer binding sites are indicated by the red arrows on the different genomic DNA templates: **Figure 6-7A:** The positive control, where the luciferase ORF is intact and integrated into the  $\beta\alpha$  tubulin array. **Figure 6-7B:** The  $\beta\alpha$  IR-LUCIFE-Act IR-BLE-Cal IR-FERASE- $\alpha\beta$  construct where recombination was not allowed to take place, due to continuous selection of phleomycin. **Figure 6-7C:** The expected product, if recombination occurred between the homologous sequences of the luciferase fragments (FE; shaded area) after removal of selection. **Figure 6-7D:** The gel shows the PCR amplification products using the Recombine Check For and Rev primers and using genomic DNA from; **Lane 1:** luciferase positive control clone A, **Lane 2** and **Lane 3:** clones 1 and 2 passaged on phleomycin, **Lane 4:** clone 1 passaged in HMI-9 without phleomycin for three months.

The lack of detectable recombination could be due to the low amount of DNA amplified in this reaction, or because the spontaneous rate of recombination is too low, or due to the constructs inability to be recombined. To begin to address this, the two clones (1 and 2) were treated with a DNA damaging agent, to increase the probability of a DSB occurring. These two clones contained the original construct  $\beta\alpha$  IR-LUCIFE-Act IR-BLE-Cal IR-FERASE- $\alpha\beta$  IR and were passaged in medium lacking phleomycin for at least three months. A total of  $1 \times 10^8$  cells from each clone were then treated with MMS at a concentration of 0.0004 % for 72 hours, genomic DNA was extracted and the PCR was performed again using the Recombine Check For and Rev primers. Again, this test did not reveal any recombination events, since the resulting PCR products obtained were approximately 2.0 kb in size (Figure 6-8). If recombination had occurred, a 476 bp PCR fragment of the size in the positive control (Figure 6-7A, Lane 1) would have been detected, but was not visible, despite overloading the gel.



**Figure 6-8: Analysing recombination of the luciferase homologous sequence by PCR-amplification after treatment with 0.0004 % MMS.**

Clone 1 and 2, which had been passaged on non-selective media for three months, were treated with 0.0004 % MMS for 72 hours. Genomic DNA was extracted from approximately  $1 \times 10^8$  cells, and PCR analysis was performed using the Recombine Check For and Rev primers.

The above result suggests that MMS-induced DNA damage does not increase the rate of HR in this assay, such that functional LUC expression is detectable. This was analysed as before, see Materials and Methods, Section 2.13. A total of  $1 \times 10^7$  cells were analysed and no detectable absorbance was measured (data not shown). It may still be the case that the assay is not sensitive enough. As a result, a large number of cells or large number of clonal populations would be required to detect LUC+ recombinants. It is also possible that more DNA damage is required, or that a different type of DNA damage may be needed. One possible solution could be the addition of an *I-SceI* endonuclease site to the construct, which would produce a DSB in the sequence between the two homologous regions when the nuclease is expressed *in vivo* (Pierce *et al.*, 1999). This would then induce DNA repair and increase the rate for HR within the assay system. Indeed, the work of Glover *et al.* (2008) suggests that the available *I-SceI* induction system induces DSBs in nearly 100 % of cells.

Another explanation is that the construct design is flawed, and is unable to undergo HR repair to lead to a functional LUC gene. The studies where the homologous regions in the gene fragments are in the same orientation with one another as direct repeat, used 1.6-3.7 kb sequence separating the gene fragments (Johnson *et al.*, 1999; Pierce *et al.*, 1999; Swoboda *et al.*, 1994). In the construct described here (Section 6.4.1) only ~1.1 kb of sequence separated the luciferase gene fragments. It is therefore possible that this sequence length is not sufficient to allow the DNA to ‘loop back’ for a recombination event to occur.

To attempt to test this, a variant of the LUC construct was generated (see Section 6.5.1 below) where the LUC fragments are orientated as inverted repeats, as it may be considered that less DNA looping is required to allow base-pairing during strand exchange.

## 6.5 An inverted repeat luciferase construct to assay homologous recombination

### 6.5.1 Generation of a luciferase construct with inverted homologous ends

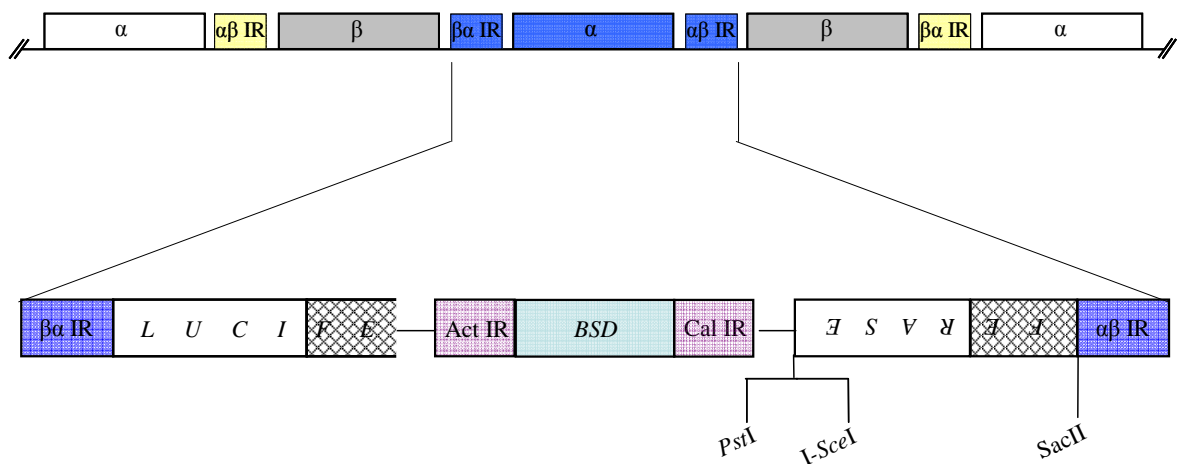
An alternative reporter construct was generated containing two inactive fragments of the luciferase reporter gene with homologous sequences orientated as inverted repeats relative to one another. In this case, these gene fragments were separated by the blasticidin antibiotic resistance cassette (*BSD*). This construct results in the homologous regions being in the same orientation when the construct forms a loop to allow recombination of the two homologous regions (see Figure 6-2B), which theoretically requires less intervening sequence between the LUC gene fragments.

This construct was generated by modifying the pBluescript:β $\alpha$ -LUCIFER-actin-BLE-calmodulin-FERASE- $\alpha\beta$  made previously (see Section 6.4.3). Firstly, the phleomycin (*BLE*) resistance ORF was replaced with the blasticidin S deaminase (*BSD*) resistance ORF. This was done by digesting the *BLE* ORF out of the construct with flanking *EcoRV* restriction sites (Figure 6-5E), and replacing it with PCR-amplified *BSD* ORF. The primers *BSD* For and Rev were used for this, which had an *EcoRV* restriction enzyme site added (see Materials and Methods, Section 2.4.1, Table 2-1). The resulting construct was checked by sequencing and was named pBluescript:β $\alpha$ -LUCIFER-actin-*BSD*-calmodulin-FERASE- $\alpha\beta$ .

In the next step, the 3' end of the luciferase gene ('*FERASE*'; in Figure 6-6) was removed by digestion with *SacII* and *PstI*. This was replaced with an inverted 3' end of the luciferase gene, which had been PCR-amplified using the "End LUC New For" primer with a *SacII* restriction site and "End LUC New Rev" primer with a *PstI* restriction site (See Materials and Methods, Section 2.4.1, Table 2-1). This PCR product was cloned using the *SacII* and *PstI* generating the construct shown in Figure 6-9.

The primer "End LUC New Rev" additionally had the target sequence of I-*SceI* nuclease incorporated meaning that the inducible I-*SceI* machinery designed by Glover *et al.* (2007), could be used to induce a DSB, forcing a recombination event to occur. The construct was

sequenced to confirm the *LUC* fragment orientation, sequence and that the *I-SceI* target site was present. Due to time constraints, this construct was not integrated or tested.



**Figure 6-9: Luciferase construct and integration into the  $\beta\alpha$  tubulin array.**

Two fragments of the luciferase gene were separated by an antibiotic resistance cassette for blasticidin (*BSD*), which was flanked by actin (Act IR) and calmodulin (Cal IR) intergenic regions. The 5' end of the luciferase gene is indicated by *LUCIFE* and the 3' of the gene is indicated by *FERASE*. The two luciferase fragments have homologous sequence (Shaded area: *FE*) and are orientated as inverted repeats.

In order to test the role of the *I-SceI* DSB induction, it will be necessary to integrate the *I-SceI* regulatory machinery into *LUC* assay-containing cells. The pHD1313 construct provides integration of the tetracycline inducible system, while the pRP-Sce construct provides the inducible *I-SceI* enzyme expression (Alibu *et al.*, 2005; Glover *et al.*, 2007). Once the luciferase construct is integrated into the  $\beta\alpha$  tubulin array (Figure 6-9), the linearised pHD1313 and pRP-Sce can then be transformed into the same cell line (Alibu *et al.*, 2005; Glover *et al.*, 2007). Again, due to time constraints, this was not carried out to completion.

## 6.6 Summary

The aim of this chapter was to design a construct which will allow *in vivo* recombination events to be directly measureable and quantifiable in *T. brucei*. An assay system similar to ones described above (Section 6-2), will allow HR to be analysed in a non-selective way.

This chapter represents the first steps in developing a novel assay for HR in *T. brucei*. Although it has left more unanswered questions than complete results, some positive data has been provided. Firstly, two luciferase-based reporter systems similar to the ones described in Section 6-2 are valid and have been generated for other model organisms. The direct repeat variant has been transformed, but appears not to spontaneously yield detectable *LUC*<sup>+</sup> recombinants. It has also been possible to express firefly luciferase in *T. brucei* wildtype bloodstream stage cells from the  $\beta\alpha$ -tubulin array, and using the substrate

d-luciferin, bioluminescence can be detected from a density of  $2.5 \times 10^5$  cells (see Figure 6-4). It has also shown that d-luciferin used in these experiments is non-toxic to *T. brucei* bloodstream stage strain 427, and that the cells can be recovered after washing with PBS and treatment with d-luciferin, and potentially could be sorted by fluorescence activated cell sorting (FACS).

Future experiments to develop this approach have been outlined, and can include the introduction of the I-SceI machinery designed by Glover *et al.* (2007) into the cells that have the  $\beta\alpha$ -LUCIFE-actin-*BSD*-calmodulin-I-SceI-ESAREF- $\alpha\beta$  integrated into the  $\alpha\beta$  tubulin array. The transformation of pHD1313 and pRP-Sce should allow the induction of a DSB within the construct, using the I-SceI site engineered between the calmodulin and 3' end fragment of *LUC*. Prior to this, it would be advantageous to transform the inverted repeat *LUC* construct into wildtype cells and ask if it spontaneously yield LUC+ recombinants.

If the assay can be developed to detect recombination, a number of important experiments can be addressed. If the assay is successful, it could be envisaged that other experiments can also be carried out to test whether or not HR is affected by, for example, a mutation in a DNA damage repair gene. DNA damage can be easily measured in different populations of varying sizes. This assay has great potential, and can be adapted to measure the effect of a number of different variables; for example, the length of homologous sequence required for recombination, the extent of sequence identity needed, whether the genomic location influence recombination, or whether specific chemicals or stresses induce recombination. It could also be adapted to provide a screen for *T. brucei* genes that act in HR. The assay could be used to ask whether life cycle variation influences HR rate.

# Chapter 7. Discussion

## 7.1 Introduction

The aim of this project was to analyse the functions of the *T. brucei* RAD51 paralogues, RAD51-4 and RAD51-6, and to examine how all four RAD51 paralogues interact in the parasite. RAD51-4 and RAD51-6 were hypothesised to have roles in homologous recombination (HR) and DNA repair, since the previously analysed RAD51 paralogues, RAD51-3 and RAD51-5, were shown to function in these pathways (Proudfoot & McCulloch, 2005a). Both *rad51-3* *-/-* and *rad51-5* *-/-* mutants display sensitivity to the DNA alkylating agent, MMS, defects in HR, slow growth rates relative to the wildtype cells, and reduction in the extent of RAD51 foci formation (Proudfoot & McCulloch, 2005a). Surprisingly, RAD51-3 but not RAD51-5 had a detectable role in antigenic variation, with the VSG switching frequency of *rad51-3* *-/-* mutants reduced compared with the wildtype and *RAD51-3* *+/-* cells while *rad51-5* *-/-* mutants had no significant reduction in VSG switching frequencies (Proudfoot & McCulloch, 2005a).

To analyse the functions of RAD51-4 and RAD51-6, homozygous mutants were generated in the *T. brucei* bloodstream stage using the Lister 427 transgenic strain 3174.2 (McCulloch *et al.*, 1997; McCulloch & Barry, 1999; Rudenko *et al.*, 1996). The phenotypes of these cell lines were investigated. To determine if the four RAD51 paralogues interacted with each other or with RAD51, yeast two-hybrid analysis, co-immunoprecipitation assays and PTP-tagging of the RAD51 paralogues were performed. The findings of these experimental approaches are discussed here. The thesis also began to develop a novel luciferase assay to measure the rate of *T. brucei* HR. This is still in the early stages and will not be discussed here (see Chapter 6, Section 6.6 for discussion).

## 7.2 Phylogeny of the *T. brucei* RAD51 paralogues

### 7.2.1 Identification of the *T. brucei* RAD51 paralogues

One of the aims of the study of RAD51-3 and RAD51-5 by Proudfoot & McCulloch (2005a) was to attempt to define RAD51-independent recombination pathways in *T. brucei*. *T. brucei rad51* *-/-* mutant cell lines showed that DNA repair, HR and antigenic variation are significantly reduced compared to the wildtype, indicating an important role for RAD51-mediated HR in these processes (Conway *et al.*, 2002b; McCulloch & Barry, 1999). Nevertheless, none of these processes were obliterated in the *rad51* *-/-* mutant cells (Conway *et al.*, 2002b; McCulloch & Barry, 1999), leading to the conclusion that alternate pathways, independent of RAD51 operate. In the absence of RAD51, recombination still

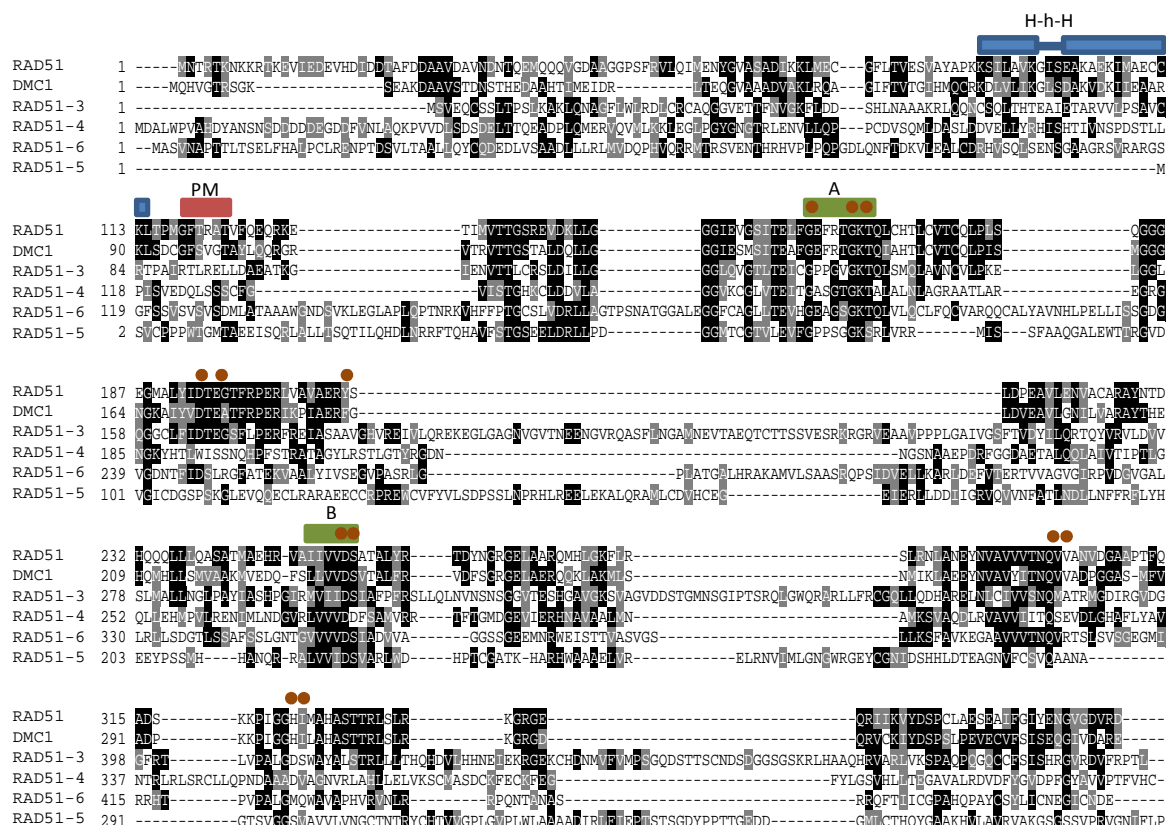
occurs by HR, though MMEJ reactions have also been revealed (Burton *et al.*, 2007; Conway *et al.*, 2002b; Glover *et al.*, 2008). However, the factors that catalyse either of these processes in *T. brucei* are unknown. To this end, Proudfoot & McCulloch (2005a) performed BLAST search analysis of the *T. brucei* genome with the amino acid sequence of *S. cerevisiae* Rad51, *T. brucei* RAD51 and *E. coli* RecA polypeptides. From this analysis five RAD51-related genes were identified (Proudfoot, PhD thesis). Searches using the amino acid sequence of *T. brucei* RAD51 revealed three distantly related, highly divergent proteins, named at the time RAD57, DMC1B and RECA2; these correspond to the current nomenclature of RAD51-3, RAD51-4 and RAD51-6 respectively. This search also revealed a gene encoding a protein more highly related to RAD51, which was shown to be an ortholog of DMC1 in other eukaryotes (Bishop *et al.*, 1992; Proudfoot & McCulloch, 2005b). Finally, another distantly related RAD51-like protein encoded in the *T. brucei* genome named RECA, corresponding to RAD51-5, was identified with a BLAST search using *E. coli* RecA amino acid sequence (Proudfoot, PhD thesis, 2005). Though the four distantly related RAD51-like proteins appear to be RAD51 paralogues (see below), a motivation for this work was that it was conceivable that the RAD51-like genes found in *T. brucei* may retain strand exchange activity. It should be noted that such a property has not been tested to date for any of the *T. brucei* RAD51 paralogues, and cannot be concluded from sequence homology.

An alignment of the amino acid sequences of the *T. brucei* RAD51, DMC1, and the RAD51 paralogues was performed (Proudfoot & McCulloch, 2005a). Figure 7-1 shows this alignment and highlights the considerable conservation of the two putative recombinases, RAD51 and the meiosis-specific DMC1, compared with the greater divergence of the RAD51 paralogues. In the core of the sequence of RAD51 and DMC1 lies the putative nucleotide/ATP-binding domains: the Walker A and Walker B motifs (Walker *et al.*, 1982). The Walker A and B motifs appear highly conserved in all four *T. brucei* paralogues, with the exception of a serine (S) residue to aspartic acid (D) mutation in the B motif of RAD51-4 (D277), suggesting they may be able to bind/hydrolyse ATP. However, any such activity and its influence on their functions can only be tested by mutagenesis or characterising the purified proteins, neither of which has been done to date. The rest of the sequence alignment suggests that sequence outwith the Walker A and B boxes appears to be less conserved between the putative recombinases and the RAD51 paralogues, in particular in motifs that are conserved between *T. brucei* RAD51 and DMC1, which are found in other species. One of these motifs is a helix-hairpin-helix (H-h-H) motif, which is a non-sequence specific DNA binding site (Shao & Grishin, 2000).



Another is a putative polymerization motif (PM), which is thought to be important in the assembly of the Rad51 nucleoprotein filament on the DNA strand (Shin *et al.*, 2003).

The Walker A motif binds ATP and the Walker B motif hydrolyses the ATP which releases the chemical energy, driving the strand exchange reaction (Story *et al.*, 1993). The roles if any, of these highly conserved motifs appear to vary in different eukaryotic Rad51 paralogues (Johnson & Symington, 1995; O'Regan *et al.*, 2001; Wiese *et al.*, 2002). The nucleotide binding and ATP-hydrolysis functions may also have important roles in the formation of mammalian Rad51 paralogue complexes, since it was found that mutation of a highly conserved lysine residue (K113) in the Walker A motif of XRCC3 was required for Rad51C-XRCC3 (CX3) complex formation (Yamada *et al.*, 2004). While it was observed that a functional Walker B motif of Rad51D was required for interaction of Rad51D with XRCC2 and Rad51C (Wiese *et al.*, 2006). The complexes Rad51B-Rad51C and Rad51D-XRCC2 have also been shown to bind to ssDNA and dsDNA, as well as having DNA-stimulated ATPase activity (Braybrooke *et al.*, 2003; Lio *et al.*, 2003; Shim *et al.*, 2004; Sigurdsson *et al.*, 2001). The CX3 complex was found to bind to ssDNA (Kurumizaka *et al.*, 2001) as well as to Holliday junctions (Liu *et al.*, 2004; Liu *et al.*, 2007). It was observed that purified RAD51D possesses DNA-stimulated ATPase activity and binds to ssDNA but not dsDNA (Braybrooke *et al.*, 2000).



**Figure 7-1: An alignment of *T. brucei* RAD51, DMC1 and the RAD51 paralogues.**

A comparison of the amino acid sequences of RAD51, DMC1 and the four RAD51 paralogues (RAD51-3, RAD51-4, RAD51-5, RAD51-6) was performed using Clustal W. Residues at a particular position with ≥ 30 % sequence identity are in black and the conserved residues are in grey between the RAD51, DMC1 and the RAD51 paralogues. The conserved Walker A and B motifs are indicated. The helix-hairpin-helix (H-h-H), putative polymerization motif (PM) and structurally conserved residues (brown circles) in RAD51 and DMC1 are indicated. Reprinted with the permission from Oxford Journals, Proudfoot & McCulloch, 2005a.

Proudfoot & McCulloch (2005a) attempted to assign functions to the *T. brucei* RAD51 paralogues by phylogenetic analysis and amino acid sequence comparison with the relatively well-studied mammalian Rad51 paralogues. Table 7-1 shows the sequence homology of the *T. brucei* RAD51 paralogues compared with the *H. sapiens* Rad51 paralogues, and with *T. brucei* RAD51, as well as the homology between the *H. sapiens* Rad51 and *H. sapiens* Rad51 paralogues. These data reinforce the view that while the recombinase, Rad51, is highly conserved with 60 % identity (70.8 % similarity) between *T. brucei* RAD51 and *H. sapiens* Rad51, the Rad51 paralogues have very low levels of sequence identity and similarity (ranging from 8.0 – 22.8 %, and 15.9 - 33.0 %, respectively) (Proudfoot, PhD thesis, 2005; Proudfoot & McCulloch., 2005a).

**Table 7-1: The percentage of amino acid sequence identities and similarities when the *T. brucei* RAD51 paralogues are compared with *T. brucei* RAD51 and *H. sapiens* Rad51 and Rad51 paralogues.**

The sequence identities (large font) and similarities (small font) are shown when the amino acid sequence of the *T. brucei* RAD51 paralogues (RAD51-3, RAD51-4, RAD51-5, RAD51-6) were compared with the *H. sapiens* Rad51 paralogues (XRCC2, XRCC3, Rad51B, Rad51C and Rad51D) using AlignX. Each of the *H. sapiens* and *T. brucei* RAD51 paralogues were also compared with the RAD51 amino acid sequence of that species (yellow shaded boxes). Adapted from Proudfoot, PhD thesis (2005) and Proudfoot & McCulloch (2005a).

		<i>H. sapiens</i>						<i>T. brucei</i> RAD51
		Rad51	XRCC2	XRCC3	Rad51B	Rad51C	Rad51D	
<i>T. brucei</i>	RAD51	60.1	14.3	20.3	20.2	19.5	16.2	100
		70.8	21.7	32	33.6	33.3	25.5	100
	RAD51-3	17.7	12.4	16.3	17	22.8	10.8	17
		27.6	19.4	27.8	26.4	33	20.3	25.9
	RAD51-4	15.5	10.9	17.4	15.8	15.7	16.4	15.6
		29.4	18.9	28.9	28.1	25.9	27.5	29.1
	RAD51-5	11.3	13.5	11.8	12.1	12.2	13.5	10.8
		22.8	23	21.1	23.7	24.3	22.7	25.2
	RAD51-6	16.8	8	17	14.9	15.6	11.7	15.9
		23.7	15.9	26.4	25.1	27	20.1	26.3
<i>H. sapiens</i>		100	16.1	21	24.7	22.2	15.8	
Rad51		100	24.9	33.6	37.1	33.1	27	

These data reveal the difficulty in attempting to assign orthologues between the mammalian Rad51 paralogues and the *T. brucei* RAD51 paralogues, meaning that assigning function thorough sequence analyses is difficult, and could even be impossible, due to the high diversity of the proteins. This underlines the fact that biochemical analysis including the interaction studies (see below), are needed to assign function to the *T. brucei* proteins, which may then clarify their relationship with the Rad51 paralogues from other organisms.

## 7.2.2 Varied number of Rad51 paralogues in different species

The recombinases, RecA, Rad51 and RadA have been highly conserved during evolution of the eubacteria, eukaryotes and archaea, respectively. Rad51 has been identified in plants such as *Arabidopsis thaliana* (Doutriaux et al., 1998), *Oryza sativa* (Ishibashi et al., 2006), fungi such as *Saccharomyces cerevisiae* (Ogawa et al., 1993), *Ustilago maydis* (Bennett & Holloman, 2001), *Coprinus cinereus* (Stassen et al., 1997), nematodes such as *Caenorhabditis elegans* (Takanami et al., 1998), parasitic protozoans including *Leishmania major* (McKean et al., 2001), *Plasmodium falciparum* (Bhattacharyya & Kumar, 2003), and *Trypanosoma brucei brucei* (McCulloch & Barry, 1999), and higher eukaryotes such

as *Mus musculus* (Cartwright *et al.*, 1998) and *Homo sapiens* (Benson *et al.*, 1994). Dmc1 also appears to be highly conserved, though it is restricted to eukaryotes (Bishop *et al.*, 1992; Doutriaux *et al.*, 1998).

In contrast, comparisons of Rad51 paralogues across the eukaryotic kingdom suggest these proteins are highly variable in sequence and also in number. Most multi-cellular organisms appear to have many Rad51 paralogues; for example, *H. sapiens* and *A. thaliana* each have five Rad51 paralogues (Bleuyard *et al.*, 2005; Masson *et al.*, 2001b), and *Drosophila melanogaster* has four (Abdu *et al.*, 2003). *Caenorhabditis elegans* appears to be an exception with only one Rad51 paralogue (Boulton *et al.*, 2002; Ward *et al.*, 2007). The budding yeast *S. cerevisiae* has two, the fission yeast *S. pombe* has four (Khasanov *et al.*, 2004; Lovett & Mortimer, 1987; Martin *et al.*, 2006; Tsutsui *et al.*, 2001) and *U. maydis* has one (Kmieć *et al.*, 1994; Rubin *et al.*, 1994). Beyond the eukaryotes, RADB in archaean species *Haloferax volcanii* has been shown to possess Rad51 paralogue like functions (Guy *et al.*, 2006). It is noteworthy, however, in eubacteria that no RecA paralogues have been found. Given that the activity of Rad51, when compared with RecA *in vitro*, is considerably reduced as discussed in Chapter 1 (Section 1.6.1) (Benson *et al.*, 1994; Conway *et al.*, 2004), it is conceivable that a more complex Rad51 pathway exists with the need for co-factors, such as the Rad51 paralogues.

If all these proteins are true Rad51 paralogues i.e. derived from Rad51 by gene duplication within the genome with the development of different functions to that of Rad51, it remains unclear why there is such a varied number. It may be the case that not all the Rad51 paralogues have been identified for those species with small numbers because of narrow searches or incomplete genome sequence. One finding which does not support this theory was first noted by Proudfoot (PhD thesis, 2005). It was observed that in *T. brucei* and *T. cruzi* four RAD51 paralogues could be identified by BLAST searches, whereas the highly related Kinetoplastida *Leishmania major* only has three RAD51 paralogues (Proudfoot, PhD Thesis). Table 7-2A shows an amino acid sequence comparison between the *T. brucei* and *T. cruzi* RAD51 paralogues, while Table 7-2B shows the same comparison between the *T. brucei* and *L. major* RAD51 paralogue amino acid sequences. It is noteworthy that the RAD51 paralogues are not found adjacent to each other, but are in disparate loci, suggesting they arose a long time ago.

**Table 7-2: The percentage of amino acid sequence identities when the *T. brucei* RAD51 paralogues were compared with the predicted *T. cruzi* and *L. major* RAD51 paralogues.**

The sequence identities are indicated as a percentage when the amino acid sequence of the *T. brucei* RAD51 paralogues (RAD51-3, RAD51-4, RAD51-5, RAD51-6) were compared to the predicted *T. cruzi* (A) and *L. major* RAD51 paralogues (B). The highest sequence identity for each comparison is indicated by yellow shading, and the predicted *T. cruzi* and *L. major* RAD51 paralogues were named accordingly. Adapted from Proudfoot, PhD thesis and Proudfoot & McCulloch., (2005a).

A		<i>T. cruzi</i>				
		RAD51	RAD51-3	RAD51-4	RAD51-5	RAD51-6
<i>T. brucei</i>	RAD51	79.9	18.1	17.1	12.8	16.7
	RAD51-3	16.1	49.8	14.6	12.5	12.7
	RAD51-4	14.6	14.3	39.9	15.6	13.3
	RAD51-5	11.4	12.4	12.4	34.8	13.8
	RAD51-6	15.3	16.4	13.8	17.1	39.2

B		<i>L. major</i>				
		RAD51	RAD51-3	RAD51-4	RAD51-5	RAD51-6
<i>T. brucei</i>	RAD51	77.5	9.6	15.6	-	10
	RAD51-3	16.3	23.9	13.2	-	11.4
	RAD51-4	15.5	12.7	18	-	12.6
	RAD51-5	9.8	7.6	12.9	-	8.7
	RAD51-6	17.2	15.1	13.8	-	21.9

Since the study by Proudfoot & McCulloch (2005a) many protozoan genomes have been sequenced, and since these represent much of the diversity in eukaryotes, it provides an opportunity to review and update the picture of the RAD51 paralogues. Other genomes such as those of *Giardia lamblia* (Aurrecochea *et al.*, 2009; Morrison *et al.*, 2007), *Trichomonas vaginalis* (Aurrecochea *et al.*, 2009; Carlton *et al.*, 2007), *Entamoeba histolytica* (Loftus *et al.*, 2005), *Dictyostelium discoideum* (Eichinger *et al.*, 2005), *Leishmania* (Peacock *et al.*, 2007), *Trypanosoma vivax* (Sanger Institute, unpublished), *Trypanosoma congolense* (Sanger Institute, unpublished) and *Encephalitozoon cuniculi* (Gill & Fast, 2007; Katinka *et al.*, 2001) have been sequenced. The genome sequences of Apicomplexans such as *Cryptosporidium parvum* (Abrahamsen *et al.*, 2004), *Toxoplasma gondii* (Gajria *et al.*, 2008), *Theileria annulata* (Pain *et al.*, 2005), and the Ciliophoran *Tetrahymena thermophila* (Eisen *et al.*, 2006) have also been completed. Based on this, BLAST searches with TbRAD51 and TbRAD51-3 were performed, then manually checked for Walker A and B boxes, as well as RecA-like signatures. This was done by BLASTing the NCBI database and identifying conserved RecA domains and hits to RecA/Rad51 (this was completed by McCulloch, 2009). The observed hits are detailed in Table 7-3. These data confirm that the Rad51 paralogues are highly variable in number throughout the eukaryotic kingdom. However, it is noteworthy that the Euglenozoa

especially the *Trypanosoma* have a large number of RAD51 paralogues, since most single celled eukaryotes have only 1-2 paralogues.

**Table 7-3: RAD51 paralogue gene numbers found in eukaryotic organisms.**

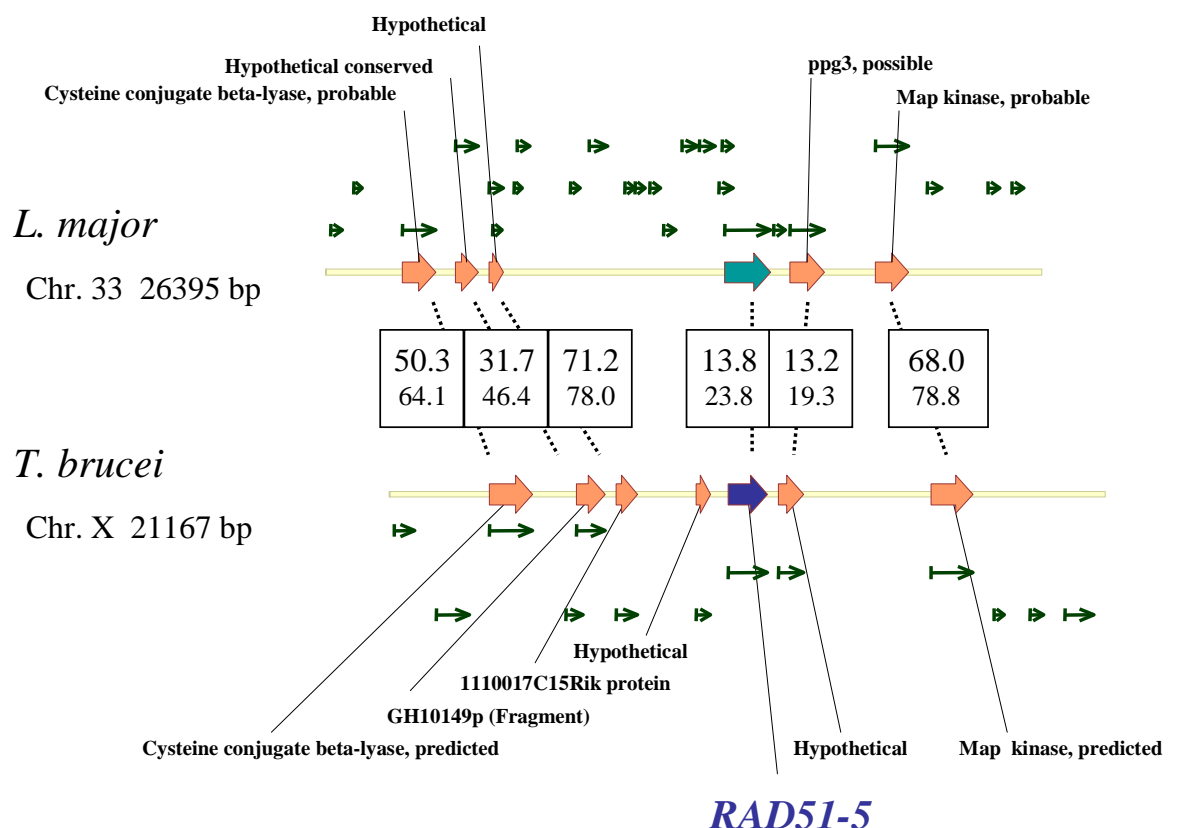
The table shows the status of the genome project for a range of different eukaryotes. It indicates the presence (+) or absence (-) of RAD51, DMC1, as well as the number of RAD51 paralogues and RecA-like genes. The RecA-like proteins are thought to act in DNA repair and HR of mitochondrial/chloroplast DNA (Cerutti *et al.*, 1992; Hasegawa *et al.*, 2004). (\* Further discussion see text; This table was provided by Dr Richard McCulloch)

Taxon	Organism	Status of genome project	RAD51	DMC1	RAD51 paralogue	RecA
Parabasala	<i>T. vaginalis</i>	7X coverage	+	+	2	-
Diplomonadida	<i>G. intestinalis</i>	11X coverage	-	2	-	-
Euglenozoa	<i>T. brucei</i>	Complete	+	+	4	-
	<i>T. vivax</i>	5X coverage	+	+	4	-
	<i>T. congolense</i>	1X coverage	+	+	4	-
	<i>T. cruzi</i>	Complete	+	+	4	-
	<i>L. major</i>	Complete	+	+	3*	-
	<i>L. braziliensis</i>	Complete	+	+	3*	-
	<i>L. infantum</i>	Complete	+	+	3*	-
Apicomplexa	<i>C. parvum</i>	Complete	+	+	2	-
	<i>T. annulatum</i>	Complete	+	+	1	-
	<i>P. falciparum</i>	Complete	+	+	1	-
	<i>T. gondii</i>	10X coverage	+	+	1	-
Ciliophora	<i>T. thermophila</i>	Complete	+	+	1	-
Viridiplantae	<i>O. sativa</i>	Complete	+	+	5	3
	<i>A. thaliana</i>	Complete	+	+	5	3
Entamoebidae	<i>E. histolytica</i>	Complete	+	+	1	-
Mycetozoa	<i>D. discoideum</i>	Complete	+	-	5	1
Microsporidia	<i>E. cuniculi</i>	Complete	+	-	1	-
Fungi	<i>S. pombe</i>	Complete	+	+	4	-
	<i>S. cerevisiae</i>	Complete	+	+	2	-
	<i>U. maydis</i>	Complete	+	-	1	-
Metazoa	<i>C. elegans</i>	Complete	+	-	1	-
	<i>D. melanogaster</i>	Complete	+	-	4	-
	<i>H. sapiens</i>	Complete	+	+	5	-
	<i>G. gallus</i>	Complete	+	+	5	-

This analysis confirms previous studies and shows that four RAD51 paralogues are present in the *Trypanosoma* species (*T. brucei*, *T. vivax*, *T. congolense* and *T. cruzi*) and that three RAD51 paralogues are found in the *Leishmania* species (*L. major*, *L. braziliensis*, and *L. infantum*; Table 7-3). This result has lead to a number of possible interpretations. The first is that all four RAD51 paralogues arose in the Kinetoplastidae by gene duplication before the evolutionary split between the *Trypanosoma* species and *Leishmania*, and undetermined differences in the biology of *Leishmania* species meant that RAD51-5 was no longer needed, whereas RAD51-5 remained important in function to the *Trypanosoma* species. The second is that RAD51-5 arose by a further gene duplication in *T. brucei*, *T.*

*vivax*, *T. congolense* and *T. cruzi* after the evolutionary split of the four *Trypanosoma* species from *Leishmania*. The third reason is that RAD51-5 was not identified in the *L. major* genome because of very low sequence identity with RAD51, and a gene encoding a RAD51-5-like protein is, in fact, present.

The completion of the genome sequence since the work of Proudfoot & McCulloch (2005a), and the sequencing of further *Leishmania* species genomes, means that this can now be answered. As a result, we believe that the first hypothesis is correct, and that *Leishmania* has lost RAD51-5. A *Leishmania* “RAD51-5-like” gene was identified by looking for orthologues of the surrounding genes of *T. brucei* RAD51-5. Figure 7-2 shows a genome synteny diagram of the two regions around the genomic location of *T. brucei* RAD51-5 and the predicted location of *L. major* “RAD51-5”. From Figure 7-2, it is clear that the surrounding genes are highly conserved showing 32-68 % sequence identity. In contrast, a putative gene containing some homology (13.8 % identity) to *T. brucei* RAD51-5 is present, and is also conserved in the other two *Leishmania* species (*L. braziliensis*, and *L. infantum*; Figure 7-3).



**Figure 7-2: Synteny diagram between the orthologues of the surrounding genes of *T. brucei* RAD51-5 and *L. major* hypothetical protein.**

The sequence identities (large font) and similarities (small font) are shown when the amino acid sequences of the surrounding proteins of *T. brucei* RAD51-5 were compared with the *L. major* hypothetical protein.

It is very questionable that the gene in this location of the *Leishmania* chromosome is a *RAD51* paralogue. The only homology is within the Walker B box motif, and the *Leishmania* “*RAD51-5?*” proteins do not possess a recognisable Walker A box, nor do they have RecA-*RAD51*-like sequence homology in the region around the Walker A box which is a characteristic of a *RAD51* paralogue. Such is the dissimilarity in these *Leishmania* “*RAD51-5?*” genes, that none were identified as orthologues in the genome annotations, in contrast to all surrounding genes. It appears that at one time, a RecA-like gene was present but has diverged to have other functions; of course, this would need to be tested experimentally. This appears compatible with the suggestion that the generation of *RAD51* paralogues in the Kinetoplastida was an ancient event, and loss of *RAD51-5* is more recent.



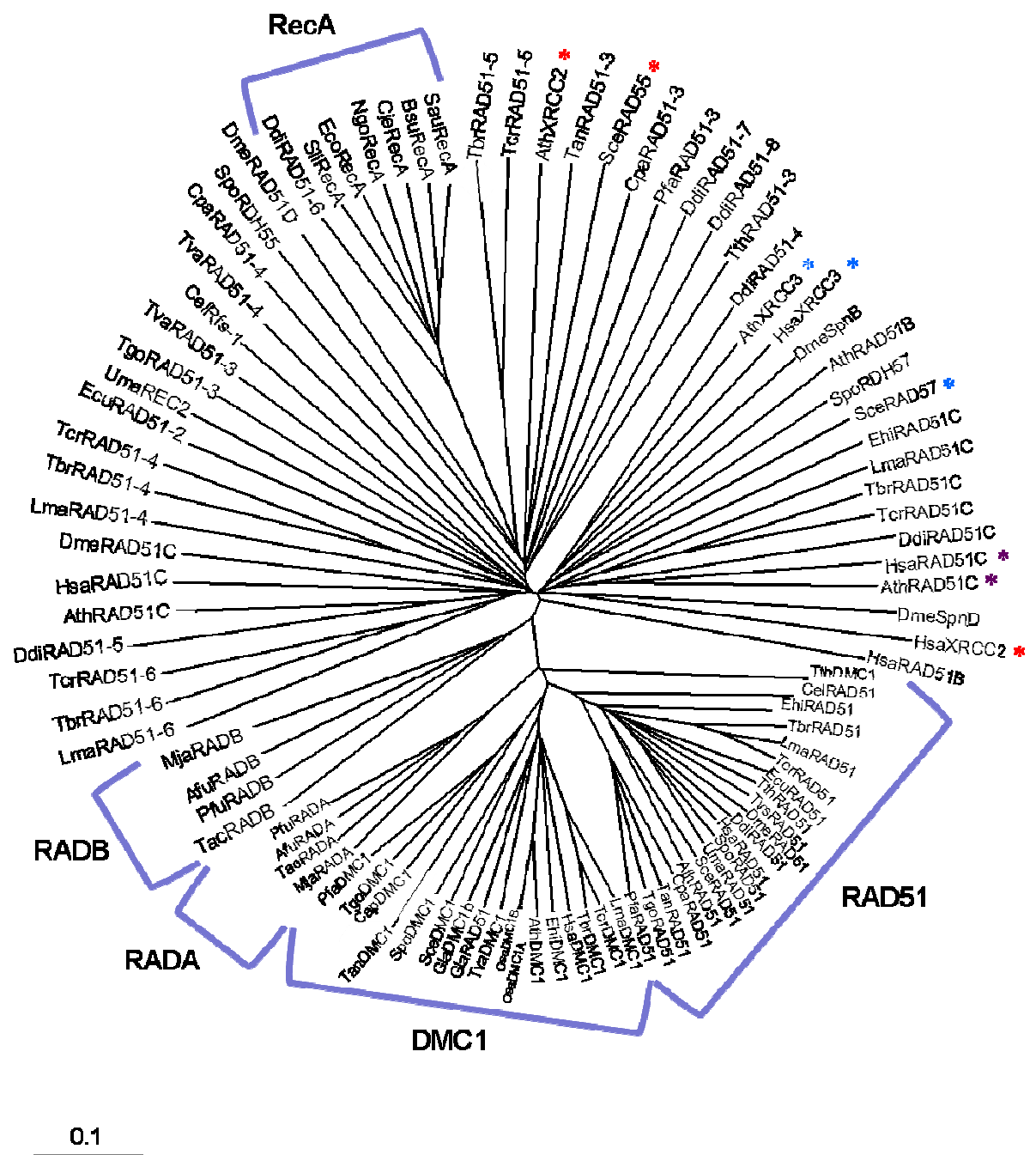
LmaRAD51-5?	1	-----	
LifRAD51-5?	1	-----	
LbrRAD51-5?	1	MQSTTASWETMSDSAARARVLRRLQSATPLQDVYGAFFSPWQHLSLSTPLGLDSLSSVAGVC	
TbrRAD51-5	1	-----	
TcrRAD51-5	1	-----	
LmaRAD51-5?	1	-----	MTMYRQWYHDE
LifRAD51-5?	1	-----	MTMCRQWYHDE
LbrRAD51-5?	61	EGVSGSDDRSGVESTIGGLPVGGLHELYGPPLSGKSWLLRRVGAAYVRRMTAYRQWYHDE	
TbrRAD51-5	1	-----	MSVCPPPPWTGMT
TcrRAD51-5	1	-----	MAVMPPPPWEALT
LmaRAD51-5?	12	LERVSKWAVETPRSTRTRYDDEAQAANKDDEGLCSEGEVADTWADSGASPVMTAMEEWDLYV	
LifRAD51-5?	12	LELASKWAVETPRSTRTRYDDEAHAKKDDEGLCSEGEVADTWADSGASPVMTAMEEWDLYV	
LbrRAD51-5?	121	LERVSKRAVEATLFTKTDNEARAAKDDEGLDSEAEADAHICANSANPSPTTAMEEWDLYV	
TbrRAD51-5	13	AEEISQRLALLTSQILQHDLNRRFTQHAFVSTGSELDRLLPDGGMCCTVLEVFEG---	
TcrRAD51-5	13	AEDVSRRLAALAASASSVLEVCRCFSREAVFSTGSAIDRLLPDGGVACGTVLEIFEG---	
A			
LmaRAD51-5?	72	CLVSGAGAGSATHOLTAPPSPSLLSPDVRSWVEELVAPFDSSTLDRQGEAFVTGPAHW	
LifRAD51-5?	72	CLVSGAGAGSATHOLTAPPSPSLLSPDVRSWVEELVAPFDSSTLDRQGEAFVTGPAHY	
LbrRAD51-5?	181	CLVRGAGADSKTHGLTASPPSPSLLSPDVRSWVEELVAPFDSSTLDRQGEAFVTGPAHW	
TbrRAD51-5	70	-----PP-----SGGKSRLVRRMTSSFAAQGALWTTRG-VDVGICD	
TcrRAD51-5	70	-----PP-----AAGKSHLVQQMVAAFAARGSVQWTAAKRSRTAVGG	
LmaRAD51-5?	132	HASTHRRQQQORDYAEQHIHFRRVVHSPNELLAFLEHLGSAALAPSAGAAFAADAAPSIVPP	
LifRAD51-5?	132	HASTHRRQQQORDYAEQHIHFRRVVHSPNELLAFLEHLGSAALAPSAGTTFADAAPSIVPP	
LbrRAD51-5?	241	YSSSH-QQQHRYAEQHIHFRRVVHSPNELLAFLECLGGGAASVPSTSTVFNTNTTYSAAP	
TbrRAD51-5	106	GSPSKGLEVQQCLRARAECC-CPRE--WCVFYVLSDPSSLNRRHLREELKALQRAML	
TcrRAD51-5	107	NSVDRQQQQQQQPHEEVESKCGPCD--WSVFLLVSDPSSVSEQHIHELKALASATA	
LmaRAD51-5?	192	VPLVASQQQPSSQLSRSSSTTTARGQKRRRSPTAVSSSSSVSPSCKLPQRTWRLQRO	
LifRAD51-5?	192	APLVASQQQPSSQLSCGSSATTARGQKRRRSPTAVSSSSSVSPFRGNFPQRTWRLQRO	
LbrRAD51-5?	300	VERVALQQQLLPSQLARGSLTAMRGQKQHRSPGTTVPSSSPPLACGSLPQCTWRLQRO	
TbrRAD51-5	163	CDVHCEG-----EERLLDDIIGRVQVNVFATLNDLNFRRFLYHEEYPSMHHANQRR	
TcrRAD51-5	165	NSVQCDG-----DADSLPVMCLKIKFVPFKSPNDLVVFFRLSRVEYFESCHHANRRR	
B			
LmaRAD51-5?	252	RLLLLDGLDALWLHPSLGNHCATHTGQWFAEELHRRRLRAVLLPRLSYAASNSAVVTPSYA	
LifRAD51-5?	252	RLLLLDGLDALWLHPSLGNHCATHTGQWFAEELHRRRLRAVLLPRLSYAASNSAVITPSYA	
LbrRAD51-5?	360	RLLLLDGLDALWLHPSLGNHSATHTGQWFAEELHRRQLRTVLLPRLGYAASDSTLTAPSS	
TbrRAD51-5	217	ALVVVDSVARLWDHPTCG--ATKHARHWAAALVRELRNVIMLGNGWRCEYCGNIDSHHL	
TcrRAD51-5	219	VLVVVDNVARLWEHPTCG--ATNHARHWAAALVREVRNAILIGNGWRDVGSGEVEKRR	
LmaRAD51-5?	312	TAAASPTSPYPHHTLYSTVVFVTNGCNGSSRGILNPQQLERLAGPVGGAEGWTATLPRPS	
LifRAD51-5?	312	TAAASPTSSHPYHTLYSTVVFVTNGCNGSSRGILNPQQLERLAGPVGGAEGWTATLPRPS	
LbrRAD51-5?	420	TAAAPMSPPHQLVLYSTVVFVTNGCHGSSRGILTAQQLERLAGPVGGAEGWTATLPRPS	
TbrRAD51-5	275	DTEAGNVFCS-----VQAANAGTSVGGSVAVVLVNGCTNTRYCHTIVG-----PLGVP-	
TcrRAD51-5	277	HHHAHDAVS-----GQEISV-DCRCGSAVAVLVNGCTSVYHRKLATERLTPSKPLGVP-	
LmaRAD51-5?	372	GNAVWCRAADTRCLVEPAHGLVSLPTPSSSYVHPARVSHGPGAMRQSRGMTAVGSPSN	
LifRAD51-5?	372	GNAVWCRAADTRCLVEPAHGLVSLPTPSSSYVCSARVSHGPGAMRQSSRGMAITFCPSN	
LbrRAD51-5?	480	GNAVWCRAADTRCLVEPAHGLVSLPTPSSSYVVRPAMAPHESAVYQNGRRITGTFROST	
TbrRAD51-5	323	--LWLAAADIRLFTPTST-----SGDYPTTGEDDDGMLCTHQYGAAKVLA---	
TcrRAD51-5	330	--VWLAAADVRLFVEPLCN-----ADWFLQHDRSSGDVCTQQSEHAARATMR---	
LmaRAD51-5?	432	ANSRESLVTVVKGGSCVAATWVLRDVTGGEQET	
LifRAD51-5?	432	TNSLESLVTVVKGGSRVAATWVLRDVAGGAQES	
LbrRAD51-5?	540	ANYLESLVTVVKGGSRVAATWVLRNADGEQEA	
TbrRAD51-5	370	-----VRVAKGSGSSVPRVGNIFTP-----	
TcrRAD51-5	377	-----VRIVKGGRTVGGEVTA-----	

**Figure 7-3: An alignment of *T. brucei* and *T. cruzi* RAD51-5 with the predicted “RAD51-5” amino acid sequence of *L. major*, *L. braziliensis*, and *L. infantum*.**

*Leishmania* proteins were identified by looking for orthologues of the surrounding genes of *T. brucei* RAD51-5. The alignment shows a comparison of the amino acid sequences of these genes and *T. brucei* and *T. cruzi* RAD51-5. This was performed using Clustal W. Residues with  $\geq 30\%$  sequence identity are in black and the conserved residues are in grey. The conserved Walker A and B motifs, RecA- RAD51-like sequence homology upstream of the Walker A motif (dashed line) and structurally conserved residues (brown circles) of *T. brucei* and *T. cruzi* RAD51-5 are indicated.

Phylogenetic studies on a large number of Rad51-like genes, including eubacterial, eukaryotic and archaean RAD51 homologues, orthologues and paralogues, has suggested that the Rad51/Dmc1/RadA group and the Rad51 paralogue group were likely to have been generated by gene duplication that predated the divergence of archaea and eukaryotes (Lin *et al.*, 2006). In this scenario, one copy of the Rad51/Dmc1/RadA group is retained in most species, reflecting its central function. Subsequent gene duplication events are then hypothesised to have occurred early in eukaryotic evolution and generated six major Rad51 paralogues which are retained in animals and plants: RadB, Rad51C, XRCC3, Rad51B, Rad51D, and XRCC2 (Lin *et al.*, 2006). It is postulated that the evolution of these multiple Rad51 paralogues in eukaryotes was driven by the necessity to maintain the fidelity of the genome, increasing the ability to repair damaged DNA and allowing genetic diversity without detriment to the cell. It maybe also be the case that the many Rad51 paralogues have evolved to support the Rad51 filament and to aid the function of the other paralogue complexes. It also appears that this diversification occurred in the archaea, as the eukaryotic RAD51 paralogues may have evolved from, or be related to RadB (Lin *et al.*, 2006).

One difficulty with the work of Lin *et al.* (2006) is that the phylogenetic analysis concentrated on Rad51 paralogues in vertebrates and plants excluding yeast, *Drosophila*, *C. elegans*, which the authors acknowledge show less clear orthology, but suggest have abbreviated Rad51 paralogue number due to gene loss. A phylogenetic tree incorporating sequences from these and protozoan organisms is shown in Figure 7-4 and their accession numbers are listed in Appendix 4, Table 15 and 16 (sequence alignment in ClustalX; tree drawn in Treeview; McCulloch, 2009). In this analysis, the close phylogenetic grouping of RAD51, DMC1, RecA, RadA and RadB proteins is retained, but the Rad51 paralogues display much greater sequence diversity. Some of the orthology detailed by Lin *et al.* (2006) can be seen by the grouping of HsaRAD51C with AthRAD51C (purple asterisks in Figure 7-4), and HsaXRCC3 with AthXRCC3 and potentially SceRad57 (blue asterisks in Figure 7-4). Some groups are much less clear, for example it is not obvious from Figure 7-4 that SceRad55 and AthXrcc2 are related to HsXrcc2 (red asterisks in Figure 7-4) as stated by Lin *et al.* (2006). Indeed, the level of likelihood of such a tree being accurate when grouping the Rad51 paralogue sequences is very low, as the bootstrap values were typically <50 % (not shown).



**Figure 7-4: A phylogenetic tree of the Rad51 related proteins.**

Diagram comparing the amino acid sequences of the Rad51 related proteins from prokaryotic and eukaryotic species. Rad51, Dmc1, RecA, RadA and RadB groups are highly reproducible with bootstrap values of >90 % (data not shown) and these groups are highlighted by brackets. This Figure was provided by Dr Richard McCulloch, 2009.

In light of these phylogenetic studies, defining the orthology and potential functions of the *T. brucei* proteins, RAD51-3, RAD51-4, RAD51-5, and RAD51-6, is problematic. It is unclear if these proteins arose from a very ancient gene duplication. Therefore should these genes be called RAD51-like genes instead of RAD51 paralogues? Nonetheless, many of these proteins have clear roles in HR and DNA repair, and all of the *T. brucei* proteins act in these processes (see below). Indeed, similar phenotypes have also been described for the mutants of *T. brucei* RAD51 paralogues when compared with the mammalian paralogues (discussed in Chapter 4; see below). Nevertheless, the amino acid sequences of the *T. brucei* proteins appear to have the highly conserved Walker A and Walker B motifs of RAD51 (see Figure 7-1) which are found in other species of Rad51 paralogues therefore defining them as RAD51 paralogues may be correct. Functional characterisation is required to determine the extent of the RAD51 paralogue orthology, and thereby clarify the evolutionary history of these genes.

## 7.3 Roles of the *T. brucei* RAD51 paralogues

### 7.3.1 DNA repair and homologous recombination in *T. brucei* RAD51 paralogue mutants

The phenotypic analyses described for the *rad51-4* *-/-* and *rad51-6* *-/-* mutant cell lines in Chapter 3 indicates that RAD51-4 and RAD51-6 have a role in DNA repair and HR in *T. brucei*, as both mutant cell lines exhibit increased sensitivity to DNA damage, defects in HR-driven integration of plasmid constructs and reduced ability to produce RAD51 foci (see below). This analysis in conjunction with previous work on RAD51-3 and RAD51-5 (Proudfoot & McCulloch, 2005a) confirm that four RAD51 paralogues are present in *T. brucei*. It also indicates that perhaps there are multiple pathways of HR in *T. brucei*, where RAD51 and the RAD51 paralogues play central roles (Barnes & McCulloch, 2007; Bell & McCulloch, 2003; Bell *et al.*, 2004; Conway *et al.*, 2002a; Conway *et al.*, 2002b; Hartley & McCulloch, 2008; McCulloch *et al.*, 1997; McCulloch & Barry, 1999; Proudfoot, 2005; Proudfoot & McCulloch, 2005b). The *in vitro* growth rates, as measured by population doubling rates of *rad51-4* *-/-* and *rad51-6* *-/-* mutants were significantly different from those of wildtype and heterozygote mutant cell lines. A similar growth defect was also seen for the *rad51-3* *-/-* and *rad51-5* *-/-* mutant cell lines (Chapter 4, Section 4.2; Proudfoot & McCulloch, 2005a). Although the cause of the increase in population doubling rate is unknown, it is probably due to increased cell death similar to that observed in the mammalian Rad51 paralogue mutant cell lines (Takata *et al.*, 2000; Takata *et al.*, 2001). Sensitivity to genotoxic agents was also significantly increased in the *rad51-4* *-/-* and

*rad51-6* *-/-* mutant cell lines, indicating that treatment with DNA damaging agents is more detrimental to the mutants compared with the wildtype. Again, this finding is comparable with phenotypes seen in *rad51-3* *-/-* and *rad51-5* *-/-* mutants (Proudfoot & McCulloch, 2005a). It is notable, however, that the *rad51-4* *-/-* mutant cell lines were less sensitive to phleomycin compared with the other three *RAD51* paralogue mutants (see Chapter 4, Section 4.3). The reason for this increased resistance is unknown, but it may imply that the role of RAD51-4 in the DNA repair pathway is not as crucial as the roles of other RAD51 paralogues. The transformation efficiencies of *rad51-4* *-/-* and *rad51-6* *-/-* mutants were each significantly reduced compared with the wildtype and corresponding heterozygous mutant cell lines, indicating that a lack of either RAD51-4 or RAD51-6 affects HR. These phenotypes are again comparable with HR rates in *rad51-3* *-/-* and *rad51-5* *-/-* mutants (Proudfoot & McCulloch, 2005a).

This analysis completes the work started by Proudfoot & McCulloch (2005a), and suggests that all four *RAD51* paralogue mutants have comparatively equivalent phenotypes, with two interesting exceptions. Firstly, the number of RAD51 foci formed in response to phleomycin. The resistance to phleomycin of the *rad51-4* *-/-* mutants compared to the other *RAD51* paralogue mutants appears to be comparable with RAD51 foci formation in response to DNA-damage by phleomycin (Chapter 3, Section 3.4.4; Chapter 4, Section 4.5). Secondly, VSG switching frequency of the *rad51-5* *-/-* mutant cell line was not significantly different from the wildtype and *RAD51-5* *+/-* cell lines (Proudfoot & McCulloch, 2005a).

### **7.3.2 RAD51 foci formation in *T. brucei* RAD51 paralogue mutants**

In mammalian cells, it has been observed that in response to DNA damage, the percentage of cells with detectable Rad51 foci increases in a time- and dose- dependent manner (Haaf *et al.*, 1995; Raderschall *et al.*, 1999). It was confirmed that the Rad51 relocalisation occurred at the site of DNA damage and was associated with the ssDNA that could result from a strand break (Raderschall *et al.*, 1999; Tashiro *et al.*, 2000). It has also been observed that Rad51 relocalises on the chromosomes of meiotic cells (Haaf *et al.*, 1995; Plug *et al.*, 1996; Scully *et al.*, 1997). *T. brucei* RAD51 relocalisation to distinct sub-nuclear foci in response to DNA damage is also observed (Glover *et al.*, 2008; Hartley & McCulloch, 2008; Proudfoot & McCulloch, 2005a). These studies confirm that Rad51 participates in nuclear processes which control chromosomal recombination and genome integrity.

RAD51 relocalisation into foci occurs at a reduced level in the four *T. brucei* RAD51 paralogue mutants compared with the wildtype (Chapter 4, Section 4.5). This was also found to be the case in *rad51* paralogue mutant chicken DT40 cell lines (Takata *et al.*, 2000; Takata *et al.*, 2001). It is possible that this residual RAD51 localisation in the individual RAD51 paralogue mutants is catalysed by other factors. One possibility is that other RAD51 paralogues might act, though this must be interpreted in the context of whether the other proteins' expression is compromised in such mutants, and whether or not they act in stable complexes (see below). Alternatively, different proteins may be involved; one strong candidate is BRCA2. *brca2* *-/-* mutant cell lines were also shown to have a reduced capability for RAD51 foci formation, suggesting that BRCA2 also has a function in RAD51 relocalisation (Hartley & McCulloch, 2008). It is unclear, however, if this is a complementary or distinct role. It has also been suggested that residual RAD51 foci formation may be due to a higher number of spontaneous DSBs in the mutants' genomes compared with wildtype cells. Repair of such spontaneous breaks may occur by RAD51 foci formation independent of the RAD51 paralogues. Whether the RAD51 paralogues protect against gross chromosomal rearrangements, indicative of spontaneous breaks, has not been analysed. Spontaneous chromosomal aberrations have been described in *T. brucei* *brca2* *-/-* and chicken *rad51* paralogue mutant cell lines, and could support this theory (Hartley & McCulloch, 2008; Takata *et al.*, 2001).

A small number (approximately 1-3 %) of non-treated *T. brucei* wildtype and RAD51 paralogue mutant cells contain observable RAD51 foci, though the reason for this is unknown (data not shown; Hartley & McCulloch, 2008; Proudfoot & McCulloch, 2005). These foci may be indicative of endogenous DNA damage repair sites, stalled replication forks or they may represent RAD51 storage sites. Rad51 foci have also been detected in around 9-13 % of non-irradiated *irs1* (*xrcc2* mutant) and *irs1SF* (*xrcc3* mutant) hamster cells and a comparable portion of non-irradiated wildtype hamster cells (Tarsounas *et al.*, 2004a). These authors suggest that the requirement for the mammalian Rad51 paralogues, XRCC2 and XRCC3, is by-passed when the Rad51 foci are associated with S-phase replication (Tarsounas *et al.*, 2004a). However, such a cell cycle stage specificity has not yet been examined in *T. brucei* mutant cell lines.

Surprisingly, *rad51-4* *-/-* mutant cell lines were more resistant to one DNA damaging agent, phleomycin, compared with *rad51* *-/-* and the other RAD51 paralogue mutants. This result may indicate that RAD51-4 has a less important role in the DNA repair pathway than the other proteins. One possibility, for instance, is that RAD51-4 acts as a non-essential co-factor for the other RAD51 paralogues, such as providing stabilisation of RAD51-3

(discussed in Section 7.4 below). This reduced phleomycin sensitivity is mirrored in the finding that the *rad51-4*  $-/-$  mutants show an increased ability to undergo RAD51 relocalisation to foci when the cells were treated with  $1.0 \mu\text{g}.\text{ml}^{-1}$  phleomycin compared with the other *RAD51* paralogue mutant cell lines (treated with the same concentration of drug). When each mutant cell line was treated with  $1.0 \mu\text{g}.\text{ml}^{-1}$  phleomycin for 18 hours, 72.7 % of wildtype cells had observable RAD51 foci, compared with 11 % of *rad51-3*  $-/-$ , 17 % of *rad51-5*  $-/-$ , and 15.3 % of *rad51-6*  $-/-$ , while in contrast, 52.7 % of *rad51-4*  $-/-$  mutant cells had observable foci. To ensure that these results were not due to a lack of RAD51 expression in the *RAD51* paralogue mutants, western blot analysis showed that each of the *RAD51* paralogue mutant cell lines were expressing RAD51 (see Chapter 4, Section 4.5), confirming that the reduction of RAD51 foci was due to a defect in the mutant cells ability to relocalise the protein. It was also shown that *rad51-/-* mutants treated with phleomycin unsurprisingly did not produce any RAD51 foci, because RAD51 is absent in these cells.

### 7.3.3 VSG switching observed in the *RAD51* paralogue mutants

The *rad51* and *RAD51* paralogue homozygous mutants were all generated in the *T. brucei* strain 3174.2, which allowed the VSG switching frequencies and mechanisms to be analysed (McCulloch *et al.*, 1997; McCulloch & Barry, 1999). Previously, it was found that, despite RAD51-3 and RAD51-5 having very similar roles in DNA repair, HR and RAD51 localisation, only the *rad51-3*  $-/-$  mutant cells were defective in VSG switching (Proudfoot & McCulloch, 2005). This indicated that, similar to RAD51, RAD51-3 contributes to *T. brucei* antigenic variation, during which HR is critical (McCulloch & Barry, 1999; McCulloch & Barry, 1999). To further examine the role of the RAD51 paralogues in *T. brucei* antigenic variation, VSG switching frequencies of the *rad51-4*  $-/-$  and *rad51-6*  $-/-$  mutant cell lines were quantified.

The *rad51-4*  $-/-$  mutant cells (X $-/-$ , Y $-/-$ ) had VSG switching frequencies of 4.06 ( $\pm$  1.57) and 5.64 ( $\pm$  1.8) switched variants per  $10^7$  cells per generation, respectively. The *rad51-6*  $-/-$  mutant cells (X $-/-$ , Y $-/-$ ) had VSG switching frequencies of 2.7 ( $\pm$  1.5) and 4.9 ( $\pm$  3.2) switched variants per  $10^7$  cells per generation, respectively. The average VSG switching frequency of the *RAD51-4* and *RAD51-6* heterozygote mutant cell lines was 12.6 switched variants per  $10^7$  cells per generation, which was ~50 % higher than wildtype cells, which had a frequency of 8.36 switched variants per  $10^7$  cells per generation (See Chapter 3, Section 3.4.6 and Section 3.7.5). The reason for this apparent increase in VSG switching

frequency of the heterozygote mutants compared with the wildtype cells is unknown. However, the VSG switching frequency of the wildtype strain 3174.2 has been shown by others to range from ~2-14 switched variants per  $10^7$  cells per generation in different experiments (Conway *et al.*, 2002a; Hartley & McCulloch, 2008; McCulloch *et al.*, 1997; McCulloch & Barry, 1999; Proudfoot & McCulloch, 2005a; Proudfoot & McCulloch, 2005b). The differences of the VSG switching frequency observed between the wildtype and heterozygote cell lines may therefore reflect the considerable variation in the VSG switching process, or in this assay. Irrespective of this, the analysis showed that the *rad51-4* *-/-* and *rad51-6* *-/-* mutants had an impaired ability to undergo VSG switching, as the frequencies with which switching variants arose in the homozygous cells were significantly lower than that of the heterozygous cell lines from which they were derived.

Compiling all of the available VSG switching data generated by this assay suggests that each of the *RAD51* paralogue mutants have, to some extent, an impaired ability to undergo antigenic variation compared with wildtype cells, as the average VSG switching frequencies for wildtype was 7.18 ( $\pm$  1.32), *rad51* *-/-* was 0.44 ( $\pm$  0.33), *rad51-3* *-/-* was 1.07 ( $\pm$  0.69), *rad51-4* *-/-* was 4.85 ( $\pm$  1.55), *rad51-5* *-/-* was 4.47 ( $\pm$  3.17) and *rad51-6* *-/-* was 3.79 ( $\pm$  2.05) switched variants per  $10^7$  cells per generation (Chapter 4, Section 4.6). These data suggest that *rad51* *-/-* and *rad51-3* *-/-* mutants have a lower mean VSG switching frequency compared with the three other *RAD51* homozygous mutants (*rad51-4* *-/-*, *rad51-5* *-/-* and *rad51-6* *-/-*), implying that RAD51 and RAD51-3 are central to VSG switching while the other paralogues may have only secondary roles. For RAD51-5, it was concluded that it does not have a role in VSG switching (Proudfoot & McCulloch, 2005a). Similarly to RAD51-5, MRE11 was shown not to affect VSG switching, even though it has clear roles in DNA repair and HR (Bell & McCulloch, 2003). The authors suggested that MRE11 and RAD51-5 may act in early or later stages of HR, which may be circumvented in the VSG switching event, although we still have no insight into the mechanism.

Hartley & McCulloch (2008) also suggested that general DNA repair and HR have different requirements to the *T. brucei* antigenic variation pathway during their study on the *BRCA2* gene. The human ortholog of *T. brucei*'s *BRCA2* was shown to interact with Rad51 via non-identical BRC repeats (Lo *et al.*, 2003; Pellegrini *et al.*, 2002). *T. brucei* *BRCA2* has an unusually large set of BRC repeats compared with other eukaryotes, containing 14 identical BRC repeats and one non-identical, C terminal repeat (Hartley & McCulloch, 2008). Phenotypic analyses showed that *T. brucei* *brca2* *-/-* mutant cell lines are defective in DNA repair, HR and antigenic variation (Hartley & McCulloch, 2008). A truncated *BRCA2* was generated in *T. brucei* cells by deleting the 14 identical repeats but



retained the divergent BRC repeat (Hartley & McCulloch, 2008). The cell line expressing this truncated BRCA2 polypeptide was impaired in DNA repair, HR and RAD51 foci formation (Hartley & McCulloch, 2008). However, these defects in the repair pathways were not reflected in the VSG switching rate (Hartley & McCulloch, 2008), which was comparable to the wildtype. These data imply that the one non-identical BRC repeat alone is unable to act with RAD51 in the DNA repair pathways of *T. brucei*, and an undefined number of distinct BRC repeats are needed for a fully functional BRCA2. On the other hand, loss of the 14 identical BRC repeats in the *BRCA2* truncated cell line did not affect VSG switching (Hartley & McCulloch, 2008), implying that the interaction between RAD51 and the BRC repeats of BRCA2 is not required for antigenic variation. As *brca2*  $\Delta$  mutant cells are defective in VSG switching, BRCA2's role in antigenic variation may be independent of RAD51 interaction. Alternatively, the HR pathway directed by RAD51 for VSG switching is distinct in its requirements compared with those for general repair. For example, the length of sequence homology needed might be quite different.

Taken as a whole, these data suggest that VSG switching is mediated by a number of the *T. brucei* DNA repair proteins, including full length BRCA2, RAD51 and at least some of the RAD51 paralogues, (RAD51-3, RAD51-4, RAD51-6; Proudfoot & McCulloch, 2005a; see Chapter 4, Section 4.6), although it is unclear how central the role for RAD51-4 and RAD51-6 are in the process. In contrast, other proteins that act in HR and DNA repair, such as RAD51-5 and MRE11, either do not have a role in antigenic variation or contribute indirectly meaning that no affect on VSG switching can be detected by the available assay (Bell & McCulloch, 2003; Proudfoot & McCulloch, 2005a). Nonetheless, it appears that *T. brucei* utilises multiple pathways of HR and it remains to be seen how many of these HR/DNA repair reactions contribute to VSG switching (Conway *et al.*, 2002b; McCulloch & Barry, 1999). As yet, all of the factors involved in each pathway are unknown.

## 7.4 Interactions between the RAD51 paralogues

Results of the phenotypic analyses of each of the four *T. brucei* RAD51 paralogue mutants are highly similar with the two exceptions of less sensitivity to phleomycin in *rad51-4*  $\Delta$  and a putative role for RAD51-3 in VSG switching. This may suggest either that these paralogues are involved at separate steps in a HR pathway, or are part of a protein complex, or complexes, involved in the pathway. There are many lines of evidence in other eukaryotes to support the hypothesis that the Rad51 paralogues act as part of multi-protein complexes. Studies with the *S. cerevisiae* Rad55 and Rad57 paralogues confirmed that both contribute as a complex to the formation of the Rad51 nucleoprotein filament in the

DNA repair pathway (see Figure 1-13). For example, it has been shown that the two paralogues interact with one another, and that Rad55 interacts with Rad51 (Hays *et al.*, 1995). Another study showed that the Rad55-Rad57 heterodimer functions with RPA (DNA binding protein) to promote Rad51 pairing on ssDNA in the early step of strand exchange (Sung, 1997). Finally, the sensitivity of *rad55* or *rad57* mutants to ionizing radiation was partially suppressed by the overexpression of RAD51 (Johnson & Symington, 1995). Double mutants of *rad55* and *rad57* show similar and often equivalent phenotypes to that of the single *rad55* or *rad57* mutants, indicating an epistatic relationship between the two genes (Johnson & Symington, 1995).

Studies with the five mammalian Rad51 paralogues confirm formation of multi-protein complexes, some involving Rad51, but here these appear to have roles in many steps of the Rad51-dependent DNA repair pathway. Two main studies suggested that the mammalian Rad51 paralogues form two distinct complexes, Rad51B-Rad51C-Rad51D-XRCC2 (BCDX2) and Rad51C-XRCC3 (CX3; see Figure 1-14), with the transient binding of Rad51 (Liu *et al.*, 2002b; Masson *et al.*, 2001b). Biochemical analyses of purified protein complexes have shown that both BCDX2 and CX3 complexes bind to ssDNA and cause DNA aggregation (Masson *et al.*, 2001a; Masson *et al.*, 2001b). As a result it was proposed that the two complexes have similar roles to that of the *S. cerevisiae* heterodimer complex Rad55/Rad57, acting in the early stages of strand exchange through their ssDNA-binding properties (Masson *et al.*, 2001a; Masson *et al.*, 2001b), perhaps stabilising the nucleoprotein filament. Interestingly, the competitive effect of RPA against Rad51 binding to ssDNA was shown to be alleviated by the Rad51B-Rad51C sub-complex (Sigurdsson *et al.*, 2001), and this complex was also shown to interact weakly with Rad51 (Liu *et al.*, 2003). The BCDX2 complex also binds to nicked duplex DNA, which perhaps implies a role in the repair of DNA (Masson *et al.*, 2001b). Analysis of purified Rad51D showed the protein has DNA binding and ATP-hydrolytic properties (Braybrooke *et al.*, 2000), and further studies indicated that Rad51D is associated with telomeres independent of the other two complex proteins, XRCC2 and Rad51C (Tarsounas *et al.*, 2004b). From this it was proposed that Rad51D has telomere protective activities, perhaps preventing telomeric attrition and fusion (Tarsounas *et al.*, 2004b). Finally, it has been proposed that Holliday junction resolvase activity is associated with the CX3 complex and that Rad51C is localised as foci in the late stages of meiotic recombination, during prophase I: pachytene and diplotene, in mouse spermatocytes (Liu *et al.*, 2007). This localisation is distinct from the Rad51 localisation, which is mainly observed as foci in the early stages of prophase I: leptotene and zygotene (Liu *et al.*, 2007). This finding, which suggests a post-synaptic role for CX3, appears to contradict the study of Takata *et al.* (2001), where Rad51 foci

formation was shown to be affected by *rad51C* and *xrcc3* mutation which would imply a function at the early step of HR. However, Liu *et al.* (2007) noted their data cannot rule out the possibility of CX3 having dual roles in the early and late stages of HR, or possibly having distinct functions in mitotic and meiotic recombination.

In summary, despite the lack of clarity in the precise roles for the yeast and mammalian RAD51 paralogues in HR, there is considerable evidence for stable complexes and in mammals, potentially multiple distinct complexes. This project sought to ascertain whether this was also true in *T. brucei*, which is highly diverged from yeast and mammals. Like the yeast and mammalian Rad51 paralogues, *T. brucei* paralogue self-interactions were not observed by yeast-two hybrid analysis. In contrast, *T. brucei* RAD51 was found to interact with itself. It seems likely that this is consistent with functional and bioinformatic prediction. RAD51 (and DMC1) act through the formation of a nucleoprotein filament, mediated by a polymerisation motif (Shin *et al.*, 2003). This motif is not recognisable in any *T. brucei* RAD51 paralogues, and has not been noted in any Rad51 paralogues from other organisms. It appears, therefore, that the paralogues interact in discrete complexes. The study of the *T. brucei* RAD51 paralogues indicated that the proteins do interact with one another: RAD51-3 and RAD51-4 interaction was confirmed by the yeast two-hybrid system, while RAD51-3 and RAD51-6 were also shown to interact, though somewhat more weakly than RAD51-3 and RAD51-4. It should be noted, that the strength of the observed interaction in the yeast two-hybrid system may not relate to true physiological and biochemical interaction. The relative stability of the interactions *in vitro* does not always correlate with their stability *in vivo*. For example, the yeast two-hybrid interaction between Rad55 and Rad51 is slightly stronger than that of Rad55 and Rad57 (Hays *et al.*, 1995; Johnson & Symington, 1995). However, experiments such as co-immunoprecipitation suggest the opposite occurs *in vivo*, with a very stable interaction between Rad55 and Rad57 (Sung, 1997). Nevertheless, the interaction between RAD51-3 and RAD51-4 was confirmed by co-immunoprecipitation from *T. brucei* cell extracts. Yeast two-hybrid analysis could not be carried out using RAD51-4 as the bait protein due to auto-induction by RAD51-4. As mentioned above, RAD51-3 and RAD51-6 were shown to interact, although the opposite was not the case and no interaction was observed between LexA-RAD51-6 and V5-NLS-B42-RAD51-3. It may be the case that this interaction is direction dependent in yeast two-hybrid analysis. There was no observable interaction between RAD51-5 and any other RAD51 paralogue in yeast two-hybrid analysis. The lack of specificity of the anti-RAD51-5 and anti-RAD51-6 anti-sera and preliminary PTP tagging purification of these proteins did not yield any information.

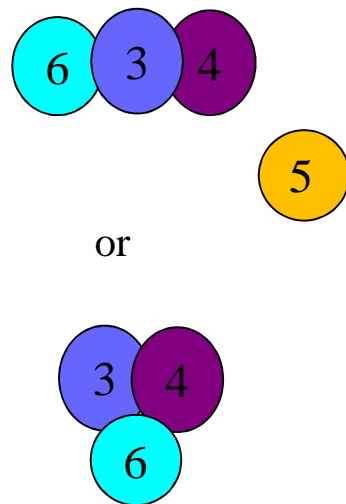
Interestingly, analysis of protein expression levels showed that in *T. brucei rad51-3*  $-/-$  mutants, RAD51-4 was undetectable. In contrast, RAD51-3 was readily detectable in *rad51-4*  $-/-$  protein extracts. This appears to suggest that the stability of RAD51-4 is dependent on the presence of RAD51-3, but not the reverse. A similar finding was shown in human cell lines: when Rad51C was depleted in HeLa cells by RNA interference, the level of XRCC3 protein was reduced, suggesting that XRCC3 is dependent on Rad51C for its stability (Lio *et al.*, 2004). In contrast, other known interactors of Rad51C, (XRCC2 and Rad51D), were not affected in the Rad51C depleted cells (Lio *et al.*, 2004). Both Rad51B and Rad51 protein levels appeared to partly reduce for day two following transfection with Rad51C siRNA, but recovered thereafter the 48 hours (Lio *et al.*, 2004). The explanation for these findings appears to lie in the existence of stable protein-protein interaction, and presumably in the absence of a binding partner, the remaining protein is targeted for destruction, either because it folds incorrectly or is trafficked aberrantly (Lehninger *et al.*, 1993). If we assume that the mammalian and *T. brucei* RAD51 paralogues have similar functions, we could postulate that *T. brucei* RAD51-3 and mammalian Rad51C have equivalent roles, stabilising RAD51-4 and XRCC3, respectively. In this context, it is interesting to note that *rad51-3*  $-/-$  mutant cell lines produced less DNA damage-induced RAD51 foci than *rad51-4*  $-/-$  mutant cell lines (discussed above, Section 7.3.2). Also, *rad51-3*  $-/-$  mutant cell lines have greater defects in HR and VSG switching, perhaps indicating that *rad51-3*  $-/-$  mutant cells are more compromised than the *rad51-4*  $-/-$  mutant cells (see Chapter 4). These differences in phenotype may simply be the result of *rad51-3*  $-/-$  mutant cell lines also lacking functional RAD51-4 (and potentially the other paralogues). One way of improving the understanding of the epistatic relationship of the two proteins would be to generate *rad51-3*  $-/-$  *rad51-4*  $-/-$  double mutants and compare the phenotype to the single mutants. If growth rate, HR, DNA damage sensitivity of the double mutants are comparable to those observed in the single *rad51-3*  $-/-$  mutant, it would demonstrate that RAD51-3 is epistatic to RAD51-4. As discussed above, *rad51-3*  $-/-$  mutants produced less phleomycin induced RAD51 foci compared with the *rad51-4*  $-/-$  mutants: 11 % of the *rad51-3*  $-/-$  mutant cells were RAD51 foci positive compared with 52.7 % of the *rad51-4*  $-/-$  mutant cells (Chapter 4, Section 4.5). If we postulate that *T. brucei* RAD51-3/RAD51-4 and mammalian Rad51C/XRCC3 have equivalent roles respectively, it is striking that the pattern of RAD51 foci observed for the *T. brucei* RAD51 paralogue mutants is similar to that of the mammalian mutant cells. After 8 hours of gamma radiation, 8 % of chicken DT40 *rad51c* mutant cells had Rad51 foci, whereas approximately double that number (16 %) had Rad51 foci in the *xrcc3* mutants (Takata *et al.*, 2001). These data are consistent with the hypothesis that *rad51c* mutant cells are somewhat more compromised than the

*xrcc3* mutants (Takata *et al.*, 2001), and therefore Rad51C has a more fundamental role in DNA repair and HR.

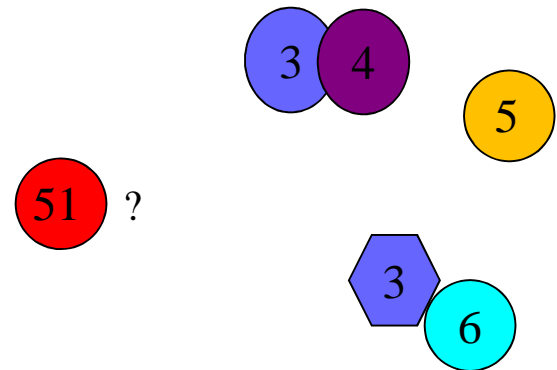
There is strong evidence that RAD51-3 and RAD51-4 interact (described as 3-4), and that RAD51-3 and RAD51-6 interact (described as 3-6) (see Figure 7-5). In contrast, we have no evidence, at least to date, that RAD51-5 interacts with the other paralogues. On the basis of these data, three possible models for interaction between the *T. brucei* RAD51 paralogues can be proposed. One is that 3-4 and 3-6 interactions form a component within a multi-protein complex (see Figure 7-5A). This could either be a linear complex with RAD51-3 at the centre or a more circular complex if 4-6 were found to interact. The other model is that 3-4 and 3-6 are two distinct heterodimeric complexes, independent of one another, perhaps similar to the mammalian BCDX2 and CX3 sub-complexes (see Figure 7-5B). It is also possible that both these models are correct (Figure 7-5), and the formation of a larger multiprotein complex can occur containing 4-3-6, while the two smaller sub-complexes 3-4, 3-6 can form and function independently. As discussed above, mammalian Rad51C was also found to be part of a larger multi-protein group (BCDX2), even though Rad51C interacts in two functional sub-complexes; Rad51B-Rad51C was shown to overcome the inhibitory affect of RPA and XRCC3-Rad51C was found to have resolvase activity. Although it is easy to see why multiple individual interactions occur simultaneously, the true function of this large complex of mammalian Rad51 paralogues (BCDX2) is unknown. The main and most plausible reason for the complex function is stability, as the complex was soluble *in vitro* while individual proteins were insoluble (Masson *et al.*, 2001b). *in vitro* assays have shown that BCDX2 binds to ssDNA, has ATP hydrolysis activity and stimulates Rad51 binding to DNA, but as yet the function of this multi-protein complex *in vivo* has not been determined.

Irrespective of which model is true for these three *T. brucei* RAD51 paralogues, the interaction, if any, between them and RAD51 and RAD51-5 is still unclear. It may be the case that RAD51 interaction is transient as observed in *S. cerevisiae* (Sung, 1997) and in mammalian cells (Sigurdsson *et al.*, 2001). It is also plausible that any interaction with Rad51 is DNA damage mediated, and to date no experiments have tested this. Yeast two-hybrid analysis failed to find evidence that RAD51-5 interacts with any of the other RAD51 paralogues, whereas there is a clear indication that the protein has a role in DNA repair, RAD51 foci formation and HR (Proudfoot & McCulloch, 2005a), confirming its definition as a RAD51 paralogue. It is intriguing that RAD51-5 was not identified in *Leishmania*, indicating perhaps that RAD51-5 has a specialised role in *T. brucei* and *T. cruzi*, which is not required in the closely related parasite.

A



B



**Figure 7-5: Models of the *T. brucei* RAD51 paralogue interactions.**

Two models for interaction of the four *T. brucei* RAD51 paralogs, RAD51-3 (3), RAD51-4 (4), RAD51-5 (5), and RAD51-6 (6) predicted from the interaction analyses data. **Figure 7-5A:** shows a multi-protein complex formed by RAD51-3, RAD51-4 and RAD51-6. **Figure 7-5B:** shows two sub-complexes formed by all the *T. brucei* RAD51 paralogs. It has not been elucidated yet where RAD51 (51) fits into this model.

## 7.5 Future goals

The results described in this thesis confirm the hypothesis that *T. brucei* possesses four RAD51 paralogues, and shows that each have roles in the HR and DNA repair pathways of the parasite. However, their precise roles in these pathways are still unknown. Analysis of cell cycle changes in *rad51*  $-/-$  and the *RAD51* paralogue homozygous mutants would determine if the slow growth phenotype is due to a cell cycle defect. Also analysis of cloning survival of the *rad51*  $-/-$  and the *RAD51* paralogue homozygous mutants would determine if slow growth was due to cell death.

Biochemical analyses of individual, purified RAD51 paralogues, as well complexes of two or three paralogues, will lead the way to a better understanding of the function of the proteins in the *T. brucei* HR pathway. Understanding their function in the pathway may also help to define the paralogues of other systems. Phylogenetic analysis of this set of proteins has proved difficult, and assigning orthology with other eukaryotes appears to be impossible due to the sequence divergence. Biochemical analyses such as DNA binding properties, HJ resolvase activity and DNA strand exchange ability, would help define the proteins in relation to the mammalian paralogues.

Analysis of the RAD51 paralogue interactions has left a number of unanswered questions. As discussed above, it is still unknown whether these proteins act in one large complex or multiple smaller complexes. Yeast two-hybrid analysis of RAD51-5 and RAD51-6 did not show interaction with any other RAD51 paralogue. Also, the purified anti-RAD51-5 and anti-RAD51-6 anti-sera did not bind specifically to the RAD51-5 and RAD51-6 proteins. As a result attempts to purify the PTP-tagged RAD51-5 and RAD51-6 were performed. Future experiments such as tagging the endogenous RAD51 and RAD51 paralogues in procyclic cell lines may yield more conclusive results. PTP tagging has been perfected in these cells. Using procyclic cells will also provide a large amount of extracted protein to improve success of the assay. As a result, this will allow mass spectrometry analysis of the precipitated proteins and will provide a comprehensive representation of what is interacting *in vivo*. Also tagging RAD51-4 and RAD51-6 with different tags such as HA or Myc tag would enable sub-cellular localisation of these proteins in the presence of DNA-damage.

One interesting finding, the co-stabilisation of RAD51-3 and RAD51-4, requires further study and confirmation. Analysing the presence of RAD51-4 by western blot in the re-expressor of RAD51-3 in the *rad51-3*  $-/-$  mutant will confirm co-stabilisation of the two

proteins. Also analysing the presence of RAD51-3 in the re-expressor of RAD51-4 in the *rad51-4*  $-/-$  mutant will also confirm co-stabilisation. Epistatic analysis of not only these two genes but also RAD51 and the other RAD51 paralogues, would indicate whether or not they act on the same or multiple pathways. Double knockout analysis, as well as overexpression studies, would lead to a better understanding of their interactions and functions within the cell.



DNA sequence (in black) and translated amino acid sequence (in blue) for RAD51-4. The positions of the primer binding sites are marked by coloured arrows.

	N Term For																
	REX For																
	Start ORF For																
	M	D	A	L	W	P	V	A	H	D	Y	A	N	S	N	S	
1	ATG	GAT	GCA	CTC	TGG	CCA	GTG	GCA	CAT	GAC	TAT	GCA	AAC	AGC	AAC	TCT	
	TAC	CTA	CGT	GAG	ACC	GGT	CAC	CGT	GTA	CTG	ATA	CGT	TTG	TCG	TTG	AGA	
	D	D	D	D	D	E	G	D	D	F	V	N	L	A	Q	K	
49	GAT	GAT	GAT	GAT	GAT	GAG	GGG	GAT	GAC	TTC	GTC	AAC	TTG	GCA	CAG	AAG	
	CTA	CTA	CTA	CTA	CTA	CTC	CCC	CTA	CTG	AAG	CAG	TTG	AAC	CGT	GTC	TTC	
	P	V	V	D	L	S	D	S	D	E	L	T	T	Q	E	A	
97	CCT	GTA	GTA	GAT	CTT	AGT	GAT	TCT	GAT	GAA	CTT	ACC	ACA	CAA	GAA	GCA	
	GGA	CAT	CAT	CTA	GAA	TCA	CTA	AGA	CTA	CTT	GAA	TGG	TGT	GTT	CTT	CGT	
	D	P	L	Q	M	E	R	V	Q	V	M	L	K	K	L	E	
145	GAT	CCA	CTC	CAA	ATG	GAG	CGT	GTT	CAA	GTT	ATG	CTG	AAG	AAG	TTG	GAG	
	CTA	GGT	GAG	GTT	TAC	CTC	GCA	CAA	GTT	CAA	TAC	GAC	TTC	TTC	AAC	CTC	
	G	L	P	G	Y	G	N	G	T	R	L	E	N	V	L	L	
193	GGC	CTT	CCA	GGT	TAT	GGG	AAC	GGC	ACG	CGT	CTT	GAA	AAC	GTG	CTA	CTG	
	CCG	GAA	GGT	CCA	ATA	CCC	TTG	CCG	TGC	GCA	GAA	CTT	TTG	CAC	GAT	GAC	
	Q	P	P	C	D	V	S	Q	M	L	D	A	S	L	D	D	
241	CAA	CCC	CCT	TGC	GAT	GTT	TCA	CAG	ATG	CTA	GAC	GCC	TCG	CTG	GAT	GAT	
	GTT	GGG	GGA	ACG	CTA	CAA	AGT	GTC	TAC	GAT	CTG	CGG	AGC	GAC	CTA	CTA	
	C Term For																
	V	E	L	L	Y	R	H	I	S	H	T	I	V	N	S	P	
289	GTG	GAG	CTT	CTC	TAC	AGA	CAC	ATA	AGC	CAC	ACC	ATC	GTT	AAC	TCT	CCC	
	CAC	CTC	GAA	GAG	ATG	TCT	GTG	TAT	TCG	GTG	TGG	TAG	CAA	TTG	AGA	GGG	
	D	S	T	L	L	P	I	S	V	E	D	Q	L	S	S	S	
337	GAT	TCC	ACG	CTT	CTT	CCC	ATT	TCC	GTT	GAA	GAC	CAA	CTC	TCC	TCC	AGT	
	CTA	AGG	TGC	GAA	GAA	GGG	TAA	AGG	CAA	CTT	CTG	GTT	GAG	AGG	AGG	TCA	
	C	F	G	V	I	S	T	G	H	K	C	L	D	D	V	L	
385	TGT	TTT	GGG	GTG	ATA	TCC	ACT	GGC	CAT	AAG	TGT	CTA	GAT	GAT	GTG	CTT	
	ACA	AAA	CCC	CAC	TAT	AGG	TGA	CCG	GTA	TTC	ACA	GAT	CTA	CTA	CAC	GAA	

529	T L A R E G R G N G K Y H T L W	ACT CTG GCA AGA GAA GGC AGG GGT AAT GGA AAG TAT CAT ACA CTT TGG	TGA GAC CGT TCT CTT CCG TCC CCA TTA CCT TTC ATA GTA TGT GAA ACC
577	I S S N Q H P F S T R A T A G Y	ATC AGC AGC AAT CAG CAT CCA TTT TCA ACC AGA GCC ACC GCC GGT TAC	TAG TCG TCG TTA GTC GTA GGT AAA AGT TGG TCT CGG TGG CGG CCA ATG
625	L R S T L G T Y R G D N N G S N	CTA CGG TCT ACA CTT GGT ACA TAT AGA GGT GAC AAC AAT GGA AGC AAT	GAT GCC AGA TGT GAA CCA TGT ATA TCT CCA CTG TTG TTA CCT TCG TTA
673	A A E P D R F G G D A E T A L Q	GCT GCT GAA CCC GAC AGG TTT GGC GGA GAT GCA GAA ACT GCG CTA CAA	CGA CGA CTT GGG CTG TCC AAA CCG CCT CTA CGT CTT TGA CGC GAT GTT
721	Q L A I V T I P T L G Q L L E H	CAG TTG GCA ATA GTA ACA ATC CCA ACG CTC GGG CAG CTT CTG GAG CAT	GTC AAC CGT TAT CAT TGT TAG GGT TGC GAG CCC GTC GAA GAC CTC GTA
769	M P V L R E N I M L N D G V R L	ATG CCT GTT CTC CGT GAG AAT ATT ATG CTG AAT GAC GGC GTA CGC CTC	TAC GGA CAA GAG GCA CTC TTA TAA TAC GAC TTA CTG CCG CAT GCG GAG
817	V V V D D F S A M V R R T F T G	GTG GTG GTG GAT GAT TTT AGC GCC ATG GTA CGT CGG ACC TTT ACT GGT	CAC CAC CAC CTA CTA AAA TCG CGG TAC CAT GCA GCC TGG AAA TGA CCA
865	M D G E V I E R H N A V A A L M	ATG GAT GGG GAG GTA ATT GAA CGG CAC AAC GCC GTT GCA GCG CTT ATG	TAC CTA CCC CTC CAT TAA CTT GCC GTG TTG CGG CAA CGT CGC GAA TAC
913	N A M K S V A Q D L R V A V V I	AAC GCA ATG AAG AGC GTT GCG CAG GAC CTT CGC GTG GCT GTT GTG ATA	TTG CGT TAC TTC TCG CAA CGC GTC CTG GAA GCG CAC CGA CAA CAC TAT
961	I T Q S E V D L G H A F L Y A V	ATT ACG CAA TCG GAA GTG GAT CTT GGT CAC GCT TTC TTG TAC GCA GTA	TAA TGC GTT AGC CTT CAC CTA GAA CCA GTG CGA AAG AAC ATG CGT CAT
1009	N T R L R L S R C L L Q P N D A	AAC ACT CGC TTG CGT TTG TCG AGA TGC TTG CTG CAA CCA AAT GAC GCA	TTG TGA GCG AAC GCA AAC AGC TCT ACG AAC GAC GTT GGT TTA CTG CGT
1057	A A D V A G N V R L A H L L E L	GCA GCA GAT GTG GCA GGT AAT GTA CGC TTG GCT CAC CTG CTG GAA CTA	CGT CGT CTA CAC CGT CCA TTA CAT GCG AAC CGA GTG GAC GAC CTT GAT
1105	V K S C M A S D C K F E C K F E	GTG AAG AGT TGC ATG GCT TCG GAC TGC AAA TTT GAG TGC AAA TTT GAG	CAC TTC TCA ACG TAC CGA AGC CTG ACG TTT AAA CTC ACG TTT AAA CTC



## Appendix 2: *RAD51-6* ORF

DNA sequence (in black) and translated amino acid sequence (in blue) for RAD51-4. The positions of the primer binding sites are marked by coloured arrows.

		<div> <div>REX For</div> <div>Start ORF For</div> </div>															
1		M	A	S	V	N	A	P	T	T	L	T	S	E	L	F	H
		ATG	GCT	AGC	GTG	AAT	GCT	CCC	ACA	ACT	CTC	ACC	AGT	GAG	CTC	TTC	CAT
		TAC	CGA	TCG	CAC	TTA	CGA	GGG	TGT	TGA	GAG	TGG	TCA	CTC	GAG	AAG	GTA
49		A	L	P	C	L	R	E	N	P	T	D	S	V	L	T	A
		GCG	CTT	CCA	TGC	CTG	CGG	GAG	AAT	CCA	ACC	GAT	AGT	GTG	CTC	ACC	GCC
		CGC	GAA	GGT	ACG	GAC	GCC	CTC	TTA	GGT	TGG	CTA	TCA	CAC	GAG	TGG	CGG
97		A	L	L	Q	Y	C	Q	D	E	D	L	V	S	A	A	D
		GCA	CTG	CTG	CAA	TAT	TGT	CAG	GAT	GAA	GAT	CTG	GTA	TCC	GCT	GCA	GAT
		CGT	GAC	GAC	GTT	ATA	ACA	GTC	CTA	CTT	CTA	GAC	CAT	AGG	CGA	CGT	CTA
145		L	L	L	R	L	M	V	D	Q	P	H	V	Q	R	R	M
		TTA	CTA	CTG	CGG	TTG	ATG	GTG	GAT	CAA	CCT	CAT	GTG	CAG	CGG	CGT	ATG
		AAT	GAT	GAC	GCC	AAC	TAC	CAC	CTA	GTT	GGA	GTA	CAC	GTC	GCC	GCA	TAC
193		T	R	S	V	E	N	T	H	R	H	V	P	L	P	Q	P
		ACA	CGT	TCT	GTC	GAG	AAT	ACA	CAT	CGA	CAT	GTG	CCG	TTG	CCT	CAA	CCT
		TGT	GCA	AGA	CAG	CTC	TTA	TGT	GTA	GCT	GTA	CAC	GGC	AAC	GGA	GTT	GGA
241		G	D	L	Q	N	F	T	D	K	V	L	E	A	L	C	D
		GGA	GAT	TTG	CAG	AAC	TTC	ACA	GAT	AAG	GTG	CTA	GAA	GCG	CTA	TGC	GAT
		CCT	CTA	AAC	GTC	TTG	AAG	TGT	CTA	TTC	CAC	GAT	CTT	CGC	GAT	ACG	CTA
																	Start ORF Rev
289		R	H	V	S	Q	L	S	E	N	S	G	A	A	G	R	S
		CGA	CAT	GTG	TCA	CAA	CTA	TCT	GAA	AAC	TCA	GGT	GCT	GCT	GGG	AGG	AGT
		GCT	GTA	CAC	AGT	GTT	GAT	AGA	CTT	TTG	AGT	CCA	CGA	CGA	CCC	TCC	TCA
337		V	R	A	R	G	S	G	F	S	S	V	S	V	S	V	S
		GTG	AGG	GCT	CGC	GGT	AGC	GGT	TTT	TCT	TCG	GTA	TCA	GTG	AGT	GTG	TCG
		CAC	TCC	CGA	GCG	CCA	TCG	CCA	AAA	AGA	AGC	CAT	AGT	CAC	TCA	CAC	AGC
385		D	M	L	A	T	A	A	A	W	G	N	D	S	V	K	L
		GAC	ATG	CTG	GCA	ACT	GCA	GCT	GCA	TGG	GGA	AAT	GAT	TCC	GTG	AAA	TTG
		CTG	TAC	GAC	CGT	TGA	CGT	CGA	CGT	ACC	CCT	TTA	CTA	AGG	CAC	TTT	AAC
433		E	G	L	A	P	L	Q	P	T	N	R	K	V	H	F	F
		GAG	GGG	CTG	GCA	CCT	TTG	CAG	CCA	ACT	AAC	CGG	AAA	GTT	CAC	TTC	TTC
		CTC	CCC	GAC	CGT	GGA	AAC	GTC	GGT	TGA	TTG	GCC	TTT	CAA	GTG	AAG	AAG
481		P	T	G	C	S	L	V	D	R	L	L	A	G	T	P	S
		CCA	ACA	GGC	TGC	TCA	CTG	GTG	GAC	CGC	CTG	CTC	GCA	GGA	ACC	CCG	TCA
		GGT	TGT	CCG	ACG	AGT	GAC	CAC	CTG	GCG	GAC	GAG	CGT	CCT	TGG	GGC	AGT

529	N	A	T	G	G	A	L	E	G	G	F	C	A	G	L	L
	AAT	GCC	ACG	GGT	GGT	GCT	CTT	GAA	GGC	GGG	TTT	TGC	GCT	GGA	CTT	CTC
	TTA	CGG	TGC	CCA	CCA	CGA	GAA	CTT	CCG	CCC	AAA	ACG	CGA	CCT	GAA	GAG
577	T	E	V	H	G	E	A	G	S	G	K	T	Q	L	V	L
	ACG	GAG	GTG	CAT	GGT	GAA	GCA	GGA	AGC	GGA	AAG	ACA	CAA	CTT	GTA	TTA
	TGC	CTC	CAC	GTA	CCA	CTT	CGT	CCT	TCG	CCT	TTC	TGT	GTT	GAA	CAT	AAT
625	Q	C	L	F	Q	C	V	A	R	Q	Q	C	A	L	Y	A
	CAG	TGC	TTA	TTC	CAA	TGC	GTT	GCT	AGG	CAA	CAG	TGT	GCT	CTG	TAT	GCT
	GTC	ACG	AAT	AAG	GTT	ACG	CAA	CGA	TCC	GTT	GTC	ACA	CGA	GAC	ATA	CGA
673	V	N	H	L	P	E	L	L	I	S	S	G	D	G	V	G
	GTG	AAT	CAC	TTA	CCC	GAG	CTG	TTG	ATC	AGC	TCA	GGC	GAT	GGT	GTT	GGT
	CAC	TTA	GTG	AAT	GGG	CTC	GAC	AAC	TAG	TCG	AGT	CCG	CTA	CCA	CAA	CCA
721	D	N	T	F	I	D	S	L	R	G	F	A	T	E	K	V
	GAC	AAC	ACG	TTT	ATT	GAT	AGT	TTG	CGT	GGC	TTT	GCA	ACT	GAA	AAG	GTG
	CTG	TTG	TGC	AAA	TAA	CTA	TCA	AAC	GCA	CCG	AAA	CGT	TGA	CTT	TTC	CAC
769	A	A	L	Y	I	V	S	E	G	V	P	A	S	R	L	G
	GCG	GCT	CTT	TAT	ATT	GTG	TCT	GAG	GGC	GTT	CCT	GCC	AGC	CGA	CTT	GGG
	CGC	CGA	GAA	ATA	TAA	CAC	AGA	CTC	CCG	CAA	GGA	CGG	TCG	GCT	GAA	CCC
817	P	L	A	T	G	A	L	H	R	A	K	A	M	V	L	S
	CCA	CTC	GCA	ACA	GGG	GCG	CTT	CAT	AGA	GCA	AAA	GCG	ATG	GTC	CTT	TCC
	GGT	GAG	CGT	TGT	CCC	CGC	GAA	GTA	TCT	CGT	TTT	CGC	TAC	CAG	GAA	AGG
865	A	A	S	R	Q	P	S	I	D	V	E	L	L	K	A	R
	GCG	GCT	TCC	CGT	CAA	CCG	TCT	ATT	GAC	GTG	GAG	CTT	TTG	AAG	GCT	CGT
	CGC	CGA	AGG	GCA	GTT	GGC	AGA	TAA	CTG	CAC	CTC	GAA	AAC	TTC	CGA	GCA
913	L	D	E	F	V	T	E	R	T	V	V	A	G	V	G	I
	CTT	GAT	GAA	TTC	GTG	ACG	GAA	CGA	ACT	GTC	GTT	GCC	GGT	GTG	GGG	ATT
	GAA	CTA	CTT	AAG	CAC	TGC	CTT	GCT	TGA	CAG	CAA	CGG	CCA	CAC	CCC	TAA
961	R	P	V	D	G	V	G	A	L	L	R	L	L	S	D	G
	CGA	CCA	GTT	GAT	GGA	GTT	GGG	GCG	TTG	TTA	CGA	CTG	CTA	TCG	GAT	GGT
	GCT	GGT	CAA	CTA	CCT	CAA	CCC	CGC	AAC	AAT	GCT	GAC	GAT	AGC	CTA	CCA
1009	T	L	S	S	A	F	S	S	L	G	N	T	G	V	V	V
	ACG	CTG	TCT	AGT	GCC	TTC	TCA	TCT	TTA	GGC	AAT	ACT	GGG	GTG	GTG	GTT
	TGC	GAC	AGA	TCA	CGG	AAG	AGT	AGA	AAT	CCG	TTA	TGA	CCC	CAC	CAC	CAA
1057	V	D	S	I	A	D	V	V	A	G	G	S	S	G	E	E
	GTA	GAT	TCC	ATT	GCA	GAT	GTA	GTG	GCT	GGT	GGA	TCT	TCG	GGC	GAG	GAG
	CAT	CTA	AGG	TAA	CGT	CTA	CAT	CAC	CGA	CCA	CCT	AGA	AGC	CCG	CTC	CTC
1105	M	N	R	W	E	I	S	T	T	V	A	S	V	G	S	L
	ATG	AAC	CGT	TGG	GAA	ATT	TCC	ACT	ACC	GTA	GCA	TCT	GTG	GGG	AGT	TTG
	TAC	TTG	GCA	ACC	CTT	TAA	AGG	TGA	TGG	CAT	CGT	AGA	CAC	CCC	TCA	AAC

End ORF For →



### **Appendix 3: Student's t-tests for the phenotypic analysis of the *rad51-4* -/- and *rad51-6* -/- mutants**

Chapter 3 describes the phenotypic analysis of the *rad51-4* -/- and *rad51-6* -/- mutants. To analysis the phenotype growth doubling times, clonal survival and IC50 values for DNA-damaging agents, transformation frequencies, and ability to form RAD51 foci were calculated. The means were calculated from three independent experiments for each assay. Unpaired, two-tailed Student's *t*-tests were carried out to compare the mean results of all the assays. These were done for the cell lines against each other to determine if there was a significant difference. A significant difference between the results is shown by a *p*-value of  $p \leq 0.05$ . Not significantly different results are shown as "-". NA is not applicable. ND is not determinable. Each cell line is described: wildtype strain 3174.2 (wt), two independent *RAD51-4/6* heterozygous mutants (X+/-, Y+/-), *rad51-4/6* homozygous mutants (X-/-, Y-/-) and re-expressor cell lines (X-/-/+, Y-/-/+)

#### **1. Statistical analysis of the doubling time of *RAD51-4* mutants.**

	wt	X+/-	Y+/-	X-/-	Y-/-	X-/-/+	Y-/-/+
wt	NA	-	-	0.0041	0.0012	-	-
X+/-	-	NA	-	0.0063	0.0028	-	-
Y+/-	-	-	NA	0.0124	0.0057	-	-
X-/-	0.0041	0.0063	0.0124	NA	-	0.0018	0.0016
Y-/-	0.0012	0.0028	0.0057	-	NA	0.0005	0.0004
X-/-/+	-	-	-	0.0018	0.0005	NA	-
Y-/-/+	-	-	-	0.0016	0.0004	-	NA

2. Statistical analysis of the clonal survival of *RAD51-4* mutants when treated with increasing concentrations of MMS (0.0001%, 0.0002%, 0.0003%, 0.0004%).

0.0001	wt	X+/-	Y+/-	X-/-	Y-/-	X-/-/+	Y-/-/+
wt	NA	-	-	-	-	-	-
X+/-	-	NA	-	-	-	-	-
Y+/-	-	-	NA	-	-	-	-
X-/-	-	-	-	NA	-	-	-
Y-/-	-	-	-	-	NA	-	-
X-/-/+	-	-	-	-	-	NA	-
Y-/-/+	-	-	-	-	-	-	NA

0.0002	wt	X+/-	Y+/-	X-/-	Y-/-	X-/-/+	Y-/-/+
wt	NA	-	-	0.0241	0.0097	-	-
X+/-	-	NA	-	0.0322	0.0130	-	-
Y+/-	-	-	NA	0.0003	0.0000	0.0298	-
X-/-	0.0241	0.0322	0.0003	NA	-	0.0027	0.0131
Y-/-	0.0097	0.0130	0.0000	-	NA	0.0004	0.0063
X-/-/+	-	-	0.0298	0.0027	0.0004	NA	-
Y-/-/+	-	-	-	0.0131	0.0063	-	NA

0.0003	wt	X+/-	Y+/-	X-/-	Y-/-	X-/-/+	Y-/-/+
wt	NA	-	-	0.0349	0.0420	-	-
X+/-	-	NA	-	0.0104	0.0256	-	-
Y+/-	-	-	NA	0.0119	0.0164	-	-
X-/-	0.0349	0.0104	0.0119	NA	-	0.0005	0.0001
Y-/-	0.0420	0.0256	0.0164	-	NA	0.0062	0.0170
X-/-/+	-	-	-	0.0005	0.0062	NA	-
Y-/-/+	-	-	-	0.0001	0.0170	-	NA

0.0004	wt	X+/-	Y+/-	X-/-	Y-/-	X-/-/+	Y-/-/+
wt	NA	-	-	0.0165	0.0165	0.0493	-
X+/-	-	NA	-	0.0342	0.0342	-	-
Y+/-	-	-	NA	0.0031	0.0031	-	-
X-/-	0.0165	0.0342	0.0031	NA	ND	0.0096	0.0096
Y-/-	0.0165	0.0342	0.0031	ND	NA	0.0096	0.0096
X-/-/+	0.0493	-	-	0.0096	0.0096	NA	-
Y-/-/+	-	-	-	0.0096	0.0096	-	NA



3. Statistical analysis of the clonal survival of *RAD51-4* mutants when treated with increasing concentrations of phleomycin ( $0.025 \mu\text{g.ml}^{-1}$ ,  $0.05 \mu\text{g.ml}^{-1}$ ,  $0.075 \mu\text{g.ml}^{-1}$ ,  $0.1 \mu\text{g.ml}^{-1}$ ).

0.1	wt	X+/-	Y+/-	X-/-	Y-/-
wt	NA	-	-	-	-
X+/-	-	NA	-	-	-
Y+/-	-	-	NA	-	-
X-/-	-	-	-	NA	ND
Y-/-	-	-	-	ND	NA

0.025	wt	X+/-	Y+/-	X-/-	Y-/-
wt	NA	-	-	0.0412	0.0306
X+/-	-	NA	-	0.0285	0.0219
Y+/-	-	-	NA	0.0347	0.0262
X-/-	0.0412	0.0285	0.0347	NA	-
Y-/-	0.0306	0.0219	0.0262	-	NA

0.05	wt	X+/-	Y+/-	X-/-	Y-/-
wt	NA	-	-	0.0139	0.0132
X+/-	-	NA	-	0.0076	0.0072
Y+/-	-	-	NA	0.0141	0.0133
X-/-	0.0139	0.0076	0.0141	NA	-
Y-/-	0.0132	0.0072	0.0133	-	NA

0.075	wt	X+/-	Y+/-	X-/-	Y-/-
wt	NA	-	-	0.0290	0.0290
X+/-	-	NA	-	0.0196	0.0196
Y+/-	-	-	NA	0.0011	0.0011
X-/-	0.0290	0.0196	0.0011	NA	ND
Y-/-	0.0290	0.0196	0.0011	ND	NA

4. Statistical analysis of the IC<sub>50</sub> values for MMS of the *RAD51-4* mutants.

	wt	X+/-	Y+/-	X-/-	Y-/-	X-/-/+	Y-/-/+
wt	NA	-	-	0.0120	0.0387	-	-
X+/-	-	NA	-	0.0063	0.0225	-	-
Y+/-	-	-	NA	0.0469	-	-	-
X-/-	0.0120	0.0063	0.0469	NA	-	-	-
Y-/-	0.0387	0.0225	-	-	NA	-	-
X-/-/+	-	-	-	-	-	NA	-
Y-/-/+	-	-	-	-	-	-	NA

5. Statistical analysis of the IC<sub>50</sub> values for phleomycin of the *RAD51-4* mutants.

	wt	X+/-	Y+/-	X-/-	Y-/-
wt	NA	-	-	0.0280	0.0499
X+/-	-	NA	-	0.0123	0.0268
Y+/-	-	-	NA	-	-
X-/-	0.0280	0.0123	-	NA	-
Y-/-	0.0499	0.0268	-	-	NA

6. Statistical analysis of the transformation efficiencies of the *RAD51-4* mutants.

	wt	X+/-	Y+/-	X-/-	Y-/-
wt	NA	-	-	0.0089	0.0047
X+/-	-	NA	-	0.0031	0.0017
Y+/-	-	-	NA	0.0373	0.0207
X-/-	0.0089	0.0031	0.0373	NA	-
Y-/-	0.0047	0.0017	0.0207	-	NA

7. Statistical analysis of the VSG switching frequencies of the *RAD51-4* mutants.

	wt	X+/-	Y+/-	X-/-	Y-/-	X-/-/+
wt	NA	0.0107	-	0.0285	-	-
X+/-	0.0107	NA	ND	0.0000	0.0001	-
Y+/-	-	ND	NA	0.0002	0.0007	-
X-/-	0.0285	0.0000	0.0002	NA	-	0.0286
Y-/-	-	0.0001	0.0007	-	NA	-
X-/-/+	-	-	-	0.0286	-	NA

8. Statistical analysis of the doubling time of *RAD51-6* mutants.

	wt	X+/-	Y+/-	X-/-	Y-/-	X-/-/+	Y-/-/+
wt	NA	-	-	0.0007	0.0004	-	-
X+/-	-	NA	-	0.0007	0.0004	-	-
Y+/-	-	-	NA	0.0008	0.0005	-	-
X-/-	0.0007	0.0007	0.0008	NA	-	0.0006	0.0009
Y-/-	0.0004	0.0004	0.0005	-	NA	0.0004	0.0006
X-/-/+	-	-	-	0.0006	0.0004	NA	-
Y-/-/+	-	-	-	0.0009	0.0006	-	NA

9. Statistical analysis of the clonal survival of *RAD51-6* mutants when treated with increasing concentrations of MMS (0.0001%, 0.0002%, 0.0003%, 0.0004%).

0.0001	wt	X+/-	Y+/-	X-/-	Y-/-	X-/-/+	Y-/-/+
wt	NA	-	-	-	0.0149	-	-
X+/-	-	NA	-	-	0.0101	-	-
Y+/-	-	-	NA	0.0434	0.0060	-	-
X-/-	-	-	0.0434	NA	-	-	-
Y-/-	0.0149	0.0101	0.0060	-	NA	0.0081	0.0127
X-/-/+	-	-	-	-	0.0081	NA	-
Y-/-/+	-	-	-	-	0.0127	-	NA

0.0002	wt	X+/-	Y+/-	X-/-	Y-/-	X-/-/+	Y-/-/+
wt	NA	-	-	0.0043	0.0132	-	-
X+/-	-	NA	-	0.0003	0.0031	-	-
Y+/-	-	-	NA	0.0000	0.0013	-	-
X-/-	0.0043	0.0003	0.0000	NA	-	0.0012	0.0126
Y-/-	0.0132	0.0031	0.0013	-	NA	0.0061	0.0314
X-/-/+	-	-	-	0.0012	0.0061	NA	-
Y-/-/+	-	-	-	0.0126	0.0314	-	NA

0.0003	wt	X+/-	Y+/-	X-/-	Y-/-	X-/-/+	Y-/-/+
wt	NA	-	-	0.0062	0.0046	-	-
X+/-	-	NA	-	0.0038	0.0029	-	-
Y+/-	-	-	NA	0.0403	0.0359	-	-
X-/-	0.0062	0.0038	0.0403	NA	-	0.0034	0.0119
Y-/-	0.0046	0.0029	0.0359	-	NA	0.0029	0.0092
X-/-/+	-	-	-	0.0034	0.0029	NA	-
Y-/-/+	-	-	-	0.0119	0.0092	-	NA

0.0004	wt	X+/-	Y+/-	X-/-	Y-/-	X-/-/+	Y-/-/+
wt	NA	-	-	-	-	-	-
X+/-	-	NA	-	-	-	-	-
Y+/-	-	-	NA	-	-	-	-
X-/-	-	-	-	NA	ND	-	0.0335
Y-/-	-	-	-	ND	NA	-	0.0335
X-/-/+	-	-	-	-	-	NA	-
Y-/-/+	-	-	-	0.0335	0.0335	-	NA

**10. Statistical analysis of the clonal survival of *RAD51-6* mutants when treated with increasing concentrations of phleomycin ( $0.025 \mu\text{g}.\text{ml}^{-1}$ ,  $0.05 \mu\text{g}.\text{ml}^{-1}$ ,  $0.075 \mu\text{g}.\text{ml}^{-1}$ ,  $0.1 \mu\text{g}.\text{ml}^{-1}$ ).**

0.025	wt	X+/-	Y+/-	X-/-	Y-/-
wt	NA	-	-	0.0002	0.0003
X+/-	-	NA	-	0.0010	0.0013
Y+/-	-	-	NA	0.0000	0.0000
X-/-	0.0002	0.0010	0.0000	NA	-
Y-/-	0.0003	0.0013	0.0000	-	NA

0.05	wt	X+/-	Y+/-	X-/-	Y-/-
wt	NA	-	-	0.0036	0.0036
X+/-	-	NA	-	0.0020	0.0020
Y+/-	-	-	NA	0.0031	0.0031
X-/-	0.0036	0.0020	0.0031	NA	ND
Y-/-	0.0036	0.0020	0.0031	ND	NA

0.075	wt	X+/-	Y+/-	X-/-	Y-/-
wt	NA	-	-	-	-
X+/-	-	NA	-	-	-
Y+/-	-	-	NA	-	-
X-/-	-	-	-	NA	ND
Y-/-	-	-	-	ND	NA

0.1	wt	X+/-	Y+/-	X-/-	Y-/-
wt	NA	-	-	-	-
X+/-	-	NA	-	ND	ND
Y+/-	-	-	NA	-	-
X-/-	-	ND	-	NA	ND
Y-/-	-	ND	-	ND	NA

**11. Statistical analysis of the IC<sub>50</sub> values for MMS of the *RAD51-6* mutants.**

	wt	X+/-	Y+/-	X-/-	Y-/-	X-/-/+	Y-/-/+
wt	NA	-	-	0.0048	0.0010	-	-
X+/-	-	NA	-	0.0038	0.0007	-	-
Y+/-	-	-	NA	0.0031	0.0006	-	-
X-/-	0.0048	0.0038	0.0031	NA	-	-	-
Y-/-	0.0010	0.0007	0.0006	-	NA	-	-
X-/-/+	-	-	-	-	-	NA	-
Y-/-/+	-	-	-	-	-	-	NA

**12. Statistical analysis of the IC<sub>50</sub> values for phleomycin of the *RAD51-6* mutants.**

	wt	X+/-	Y+/-	X-/-	Y-/-
wt	NA	-	-	0.0060	0.0108
X+/-	-	NA	-	0.0101	0.0175
Y+/-	-	-	NA	0.0022	0.0042
X-/-	0.0060	0.0101	0.0022	NA	-
Y-/-	0.0108	0.0175	0.0042	-	NA

**13. Statistical analysis of the transformation efficiencies of the *RAD51-6* mutants.**

	wt	X+/-	Y+/-	X-/-	Y-/-
wt	NA	-	-	0.0220	0.0148
X+/-	-	NA	-	0.0364	0.0265
Y+/-	-	-	NA	0.0502	0.0373
X-/-	0.0220	0.0364	0.0502	NA	-
Y-/-	0.0148	0.0265	0.0373	-	NA

**14. Statistical analysis of the VSG switching frequencies of the *RAD51-6* mutants**

	wt	X+/-	Y+/-	X-/-	Y-/-	X-/-/+
wt	NA	-	-	0.0328	-	-
X+/-	-	NA	-	0.0054	0.0487	-
Y+/-	-	-	NA	0.0011	0.0181	-
X-/-	0.0328	0.0054	0.0011	NA	-	-
Y-/-	-	0.0487	0.0181	-	NA	-
X-/-/+	-	-	-	-	-	NA

#### **Appendix 4: Accession numbers of the RAD51-related genes**

The accession numbers for the Rad51, RecA, RadA, RadB and RAD51-like genes used during the phylogenetic analysis are shown in Table 15 and 16. The sequences were obtained from *Trypanosoma brucei* (Tbr), *Leishmania major* (Lma), *Trypanosoma cruzi* (Tcr), *Entamoeba histolytica* (Ehi), *Giardia lamblia* (Gla), *Dictyostelium discoideum* (Ddi), *Toxoplasma gondii* (Tgo), *Plasmodium falciparum* (Pfa), *Cryptosporidium parvum* (Cpa), *Theileria annulata* (Tan), *Tetrahymena thermophila* (Tth), *Trichomonas vaginalis* (Tva), *Saccharomyces cerevisiae* (Sce), *Homo sapiens* (Hsa), *Arabidopsis thaliana* (Ath), *Schizosaccharomyces pombe* (Spo), *Drosophila melanogaster* (Dme), *Caenorhabditis elegans* (Cel), *Ustiligo maydis* (Uma), *Encephalitozoon cuniculi* (Ecu), *Oryza sativa* (Osa), *Escherichia coli* (Eco), *Campylobacter jejuni* (Cje), *Neisseria gonorrhoeae* (Ngo), *Bacillus subtilis* (Bsu), *Streptomyces lividans* (Sli), *Staphylococcus aureus* (Sau), *Pyrococcus furiosus* (Pfu), *Methanococcus jannaschii* (Mja), *Archaeoglobus fulgidus* (Afu), *Thermoplasma acidophilum* (Tac).

##### **15. Accession numbers of the *T. brucei* RAD51, DMC1 and the RAD51paralogues.**

TbrRAD51	Tb11.01.0360
TbrDMC1	Tb09.211.1210
TbrRAD51-4	Tb11.02.4880
TbrRAD51-3	Tb11.02.0150
TbrRAD51-5	Tb10.389.1770
TbrRAD51-6	Tb927.3.5230

## 16. Accession numbers of the RAD51-related genes

Species	Accession Number
LmaRAD51	AAC16334; LmjF28.0550
LmaDMC1	LmjF35.4890
LmaRAD51-3	LmjF33.2490
LmaRAD51-4	LmjF11.0230
LmaRAD51-6	LmjF29.0450
TcrRAD51	Tc00.1047053503801.30
TcrDMC1	Tc00.1047053506885.310
TcrRAD51-3	Tc00.1047053504153.220
TcrRAD51-4	Tc00.1047053503613.30
TcrRAD51-6	Tc00.1047053508075.20
TcrRAD51-5	Tc00.1047053510123.30
EhiRAD51	63.m00157
EhiDMC1	10.m00382
EhiRAD51-3	RAD51C; 59.m00167
GlaRAD51	XP_001709425; GL50803_13104
GlaDMC1b	XP_001710001; GL50803_13346
DdiRAD51	DDB0168161; DDB0217219
DdiRAD51-3	RAD51c; DDB0186043
DdiRAD51-4	xrcc3; DDB0185591
DdiRAD51-6	DDB0169470; mitochondrial
DdiRAD51-7	DDB0188824
DdiRAD51-8	DDB0217538
TgoRAD51	ToxoDB:59.m03654
TgoDMC1	ToxoDB:35.m00010
TgoRAD51-3	645.m00312
PfaRAD51	PlasmoDB:PF11_0087
PfaDMC1	PlasmoDB:MAL8P1.76
PfaRAD51-3	PFD0935c
CpaRAD51	CryptoDB:cgd5_410
CpaDMC1	CryptoDB:cgd7_1690
CpaRAD51-3	cgd2_4070
CpaRAD51-4	cgd6_4800
TanRAD51	TA07350
TanDMC1	TA07075
TanRAD51-3	TA12290
TthRAD51	TTERM_00142330
TthDMC1	TTHERM_00459230
TthRAD51-3	TTHERM_01143840
TvaRAD51	TVAG_204070
TvaDMC1	TVAG_155030
TvaRAD51-3	TVAG_426330
TvaRAD51-4	TVAG_144570
SceRAD51	NP_011021
SceDMC1	NP_011106
SceRAD57	NP_010287
SceRAD55	NP_010361

Species	Accession Number
HsaRAD51	Q06609
HsaDMC1	Q14565
HsaXRCC2	O43543
HsaXRCC3	O43542
HsaRAD51B	O15315
HsaRAD51D	AAC39719
HsaRAD51C	AAC39604
AthRAD51	NP_568402
AthDMC1	JC4092
AthRAD51C	NP_566040
AthRAD51B	NP_180423
AthRAD51D	NP_172254
AthXRCC2	NP_201257
AthXRCC3	NP_851202
SpoRHP51	P36601
SpoDMC1	O42634
SpoRDH55	O14129
SpoRDH57	Q9UUL2
DmeRAD51	BAA04580
DmeSpnD	NP_733200
DmeSpnB	NP_476740
DmeRAD51C	NP_610466
DmeRAD51D	NP_573302
CelRAD51	O44246 O44246_CAEEL
CelRfs-1	P34348 YK82_CAEEL
UmaRAD51	AAC61878
UmaREC2	A56244
EcuRAD51	Q8SQX0_ENCCU
EcuRAD51-2	Q8SW20_ENCCU
OsaDMC1A	AAM76793
OsaDMC1B	AAM76792
EcoRECA	P03017
CjeRECA	CAB73660
NgoRECA	CAA35247
BsuRECA	P16971
SliRECA	P48294
SauRECA	CAG40263
PfuRADA	NP_579655
MjaRADA	E64408
AfuRADA	O29269 RADA_ARCFU
TacRADA	Q9HJ68 RADA_THEAC
PfuRADB	P81415 RADB_PYRFU
AfuRADB	O28184 RADB_ARCFU
TacRADB	Q9HJD3 RADB_THEAC
MjaRADB	Q57702 RADB_METJA



## **Bibliography**

**Abbas, A. K., Lichtman, A. H. & Pober, J. S. (1994).** *Cellular and Molecular Immunology*. New York, USA: W. B. Saunders Company.

**Abdu, U., Gonzalez-Reyes, A., Ghabrial, A. & Schupbach, T. (2003).** The *Drosophila* *spn-D* gene encodes a RAD51C-like protein that is required exclusively during meiosis. *Genetics* **165** (1), 197-204.

**Abrahamsen, M. S., Templeton, T. J., Enomoto, S., Abrahante, J. E., Zhu, G., Lancto, C. A., Deng, M., Liu, C., Widmer, G., Tzipori, S., Buck, G. A., Xu, P., Bankier, A. T., Dear, P. H., Konfortov, B. A., Spriggs, H. F., Iyer, L., Anantharaman, V., Aravind, L. & Kapur, V. (2004).** Complete Genome Sequence of the Apicomplexan, *Cryptosporidium parvum*. *Science* **304** (5669), 441-445.

**Aihara, H., Ito, Y., Kurumizaka, H., Terada, T., Yokoyama, S. & Shibata, T. (1997).** An interaction between a specified surface of the C-terminal domain of RecA protein and double-stranded DNA for homologous pairing. *Journal of Molecular Biology* **274** (2), 213-221.

**Aihara, H., Ito, Y., Kurumizaka, H., Yokoyama, S. & Shibata, T. (1999).** The N-terminal domain of the human Rad51 protein binds DNA: structure and a DNA binding surface as revealed by NMR. *Journal of Molecular Biology* **290** (2), 495-504.

**Albala, J. S., Thelen, M. P., Prange, C., Fan, W., Christensen, M., Thompson, L. H. & Lennon, G. G. (1997).** Identification of a novel human Rad51 homolog, Rad51B. *Genomics* **46** (3), 476-479.

**Alexiadis, V. & Kadonaga, J. T. (2002).** Strand pairing by Rad54 and Rad51 is enhanced by chromatin. *Genes & Development* **16** (21), 2767-2771.

**Alibu, V. P., Storm, L., Haile, S., Clayton, C. & Horn, D. (2005).** A doubly inducible system for RNA interference and rapid RNAi plasmid construction in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology* **139** (1), 75-82.

**Alifano, P., Fani, R., Lio, P., Lazcano, A., Bazzicalupo, M., Carlomagno, M. S. & Bruni, C. B. (1996).** Histidine biosynthetic pathway and genes: structure, regulation, and evolution. *Microbiological Review* **60** (1), 44-69.

**Alsford, N. S., Navarro, M., Jamnadass H. R., Dunbar H., Ackroyd, M., Murphy, N.B., Gull, K. & Ersfeld, K. (2003).** The identification of circular extrachromosomal DNA in the nuclear genome of *Trypanosoma brucei*. *Molecular Microbiology* **47** (2), 277-289.

**Anoopkumar-Dukie, S., Carey, J. B., Conere, T., O'Sullivan, E., van Pelt, F. N. & Allshire, A. (2005).** Resazurin assay of radiation response in cultured cells. *The British Journal of Radiology* **78** (934), 945-947.

**Auffret, C. A. & Turner, M. J. (1981).** Variant specific antigens of *Trypanosoma brucei* exist in solution as glycoprotein dimers. *The Biochemical Journal* **193** (2), 647-650.

**Aurrecoechea, C., Brestelli, J., Brunk, B. P., Carlton, J. M., Dommer, J., Fischer, S., Gajria, B., Gao, X., Gingle, A., Grant, G., Harb, O. S., Heiges, M., Innamorato, F., Iodice, J., Kissinger, J. C., Kraemer, E., Li, W., Miller, J. A., Morrison, H. G., Nayak,**

**V., Pennington, C., Pinney, D. F., Roos, D. S., Ross, C., Stoeckert, C. J., Jr., Sullivan, S., Treatman, C. & Wang, H. (2009).** GiardiaDB and TrichDB: integrated genomic resources for the eukaryotic protist pathogens *Giardia lamblia* and *Trichomonas vaginalis*. *Nucleic Acids Research* **37** 526-530.

**Barbour, A., Dai, Q., Restrepo, B. I., Stoenner, H. G. & Frank, S. A. (2006).** Pathogen escape from host immunity by a genome program for antigenic variation. *Proceedings of the National Academy of Sciences* **13** (48), 18290-18295.

**Barbour, A. G. & Restrepo, B. I. (2000).** Antigenic variation in vector-borne pathogens. *Emerging Infectious Disease* **6** (5), 449-457.

**Barnes, R. & McCulloch, R. (2007).** *Trypanosoma brucei* homologous recombination is dependent on substrate length and homology, though displays a differential dependence on mismatch repair as substrate length decreases. *Nucleic Acids Research* **35** (10), 3478-3493.

**Barrett, M. P., Burchmore, R. J. S., Stich, A., Lazzari, J. O., Frasc, A. C., Cazzulo, J. J. & Krishna, S. (2003).** The trypanosomiasis. *The Lancet* **362** (9394), 1469-1480.

**Barry, J. & McCulloch, R. (2001).** Antigenic variation in trypanosomes: enhanced phenotypic variation in a eukaryotic parasite. *Advances in Parasitology* **49** 1-55.

**Barry, J. D. (1997).** The relative significance of mechanisms of antigenic variation in African trypanosomes. *Parasitology Today* **13** (6), 212-218.

**Barry, J. D., Marcello, L., Morrison, L. J., Read, A. F., Lythgoe, K., Jones, N., Carrington, M., Blandin, G., Bahme, U., Caler, E., Hertz-Fowler, C., Renault, H., El-Sayed, N. & Berriman, M. (2005).** What the genome sequence is revealing about trypanosome antigenic variation. *Biochemical Society Transactions* **33** (5), 986-989.

**Baumann, P., Benson, F. E. & West, S. C. (1996).** Human Rad51 protein promotes ATP-dependent homologous pairing and strand transfer reactions *in-vitro*. *Cell* **87** (4), 757-766.

**Baumann, P. & West, S. C. (1998).** Role of the human Rad51 protein in homologous recombination and Double-Stranded-Break repair. *Trends in Biochemical Sciences* **23** (7), 247-251.

**Becker, M., Aitcheson, N., Byles, E., Wickstead, B., Louis, E. & Rudenko, G. (2004).** Isolation of the repertoire of VSG expression site containing telomeres of *Trypanosoma brucei* 427 using transformation-associated recombination in yeast. *Genome Research* **14** (11), 2319-2329.

**Bell, J. S. & McCulloch, R. (2003).** Mismatch repair regulates homologous recombination, but has little influence on antigenic variation in *Trypanosoma brucei*. *Journal of Biological Chemistry* **278** (46), 45182-45188.

**Bell, J. S., Harvey, T. I., Sims, A. & McCulloch, R. (2004).** Characterization of components of the mismatch repair machinery in *Trypanosoma brucei*. *Molecular Microbiology* **51** (1), 159-173.

**Bennett, R. L. & Holloman, W. K. (2001).** A RecA homologue in *Ustilago maydis* that is distinct and evolutionarily distant from Rad51 actively promotes DNA pairing reactions in the absence of auxiliary factors. *Biochemistry* **40** (9), 2942-2953.

- Benson, F., Stasiak, A. & West, S. C. (1994).** Purification and characterization of the human Rad51 protein, an analogue of *E. coli* RecA. *The EMBO Journal* **13** (23), 5764-5771.
- Bentley, J., Diggle, C. P., Harnden, P., Knowles, M. A. & Kiltie, A. E. (2004).** DNA double strand break repair in human bladder cancer is error prone and involves microhomology-associated end-joining. *Nucleic Acids Research* **32** (17), 5249-5259.
- Beranek, D. T. (1990).** Distribution of methyl and ethyl adducts following alkylation with monofunctional alkylating agents. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **231** (1), 11-30.
- Berriman, M., Hall, N., Sheader, K., Bringaud, F., Tiwari, B., Isobe, T., Bowman, S., Corton, C., Clark, L. & Cross, G. (2002).** The architecture of variant surface glycoprotein gene expression sites in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology* **122** (2), 131-140.
- Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renauld, H., Bartholomeu, D. C., Lennard, N. J., Caler, E., Hamlin, N. E., Haas, B., Bohme, U., Hannick, L., Aslett, M. A., Shallom, J., Marcello, L., Hou, L., Wickstead, B., Alsmark, U. C. M., Arrowsmith, C., Atkin, R. J., Barron, A. J., Bringaud, F., Brooks, K., Carrington, M., Cherevach, I., Chillingworth, T.-J., Churcher, C., Clark, L. N., Corton, C. H., Cronin, A., Davies, R. M., Doggett, J., Djikeng, A., Feldblyum, T., Field, M. C., Fraser, A., Goodhead, I., Hance, Z., Harper, D., Harris, B. R., Hauser, H., Hostetler, J., Ivens, A., Jagels, K., Johnson, D., Johnson, J., Jones, K., Kerhornou, A. X., Koo, H., Larke, N., Landfear, S., Larkin, C., Leech, V., Line, A., Lord, A., MacLeod, A., Mooney, P. J., Moule, S., Martin, D. M. A., Morgan, G. W., Mungall, K., Norbertczak, H., Ormond, D., Pai, G., Peacock, C. S., Peterson, J., Quail, M. A., Rabbinowitsch, E., Rajandream, M.-A., Reitter, C., Salzberg, S. L., Sanders, M., Schobel, S., Sharp, S., Simmonds, M., Simpson, A. J., Tallon, L., Turner, C. M. R., Tait, A., Tivey, A. R., Van Aken, S., Walker, D., Wanless, D., Wang, S., White, B., White, O., Whitehead, S., Woodward, J., Wortman, J., Adams, M. D., Embley, T. M., Gull, K., Ullu, E., Barry, J. D., Fairlamb, A. H., Opperdoes, F., Barrell, B. G., Donelson, J. E., Hall, N., Fraser, C. M., Melville, S. E. & El-Sayed, N. M. (2005).** The genome of the African Trypanosome *Trypanosoma brucei*. *Science* **309** (5733), 416-422.
- Bhattacharyya, M. K. & Kumar, N. (2003).** Identification and molecular characterisation of DNA damaging agent induced expression of *Plasmodium falciparum* recombination protein PfRad51. *International Journal for Parasitology* **33** (12), 1385-1392.
- Bhattacharyya, M. K., Norris, D. E. & Kumar, N. (2004).** Molecular players of homologous recombination in protozoan parasites: implications for generating antigenic variation. *Infection, Genetics and Evolution* **4** (2), 91-98.
- Bishop, D. K., Park, D., Xu, L. & Kleckner, N. (1992).** DMC1: A meiosis-specific yeast homolog of *E. coli* recA required for recombination, synaptonemal complex formation, and cell cycle progression. *Cell* **69** (3), 439-456.
- Bishop, D. K. (1994).** RecA homologs Dmc1 and Rad51 interact to form multiple nuclear complexes prior to meiotic chromosome synapsis. *Cell* **79** (6), 1081-1092.
- Bishop, D. K., Ear, U., Bhattacharyya, A., Calderone, C., Beckett, M., Weichselbaum, R. R. & Shinohara, A. (1998).** Xrcc3 is required for assembly of Rad51 complexes *in vivo*. *Journal of Biological Chemistry* **273** (34), 21482-21488.

- Bleuyard, J.-Y., Gallego, M. E., Savigny, F. & White, C. I. (2005).** Differing requirements for the *Arabidopsis* Rad51 paralogs in meiosis and DNA repair. *The Plant Journal* **41** (4), 533-545.
- Bliss, T. M. & Lane, D. P. (1997).** Ku selectively transfers between DNA molecules with homologous ends. *Journal of Biological Chemistry* **272** (9), 5765-5773.
- Block, W. D., Yu, Y., Merkle, D., Gifford, J. L., Ding, Q., Meek, K. & Lees-Miller, S. P. (2004).** Autophosphorylation-dependent remodeling of the DNA-dependent protein kinase catalytic subunit regulates ligation of DNA ends. *Nucleic Acids Research* **32** (14), 4351-4357.
- Blum, M. L., Down, J. A., Gurnett, A. M., Carrington, M., Turner, M. J. & Wiley, D. C. (1993).** A structural motif in the variant surface glycoproteins of *Trypanosoma brucei*. *Nature* **362** (6421), 603-609.
- Blumenthal, T. (2005).** *trans*-slicing and operons. *Wormbook: the online review of C elegans biology* 1-9 ([www.wormbook.org](http://www.wormbook.org)).
- Boothroyd, J. C. & Konuniecki, R. (1995).** *Molecular approaches to parasitology*: Wiley-Liss.
- Borst, P., Rudenko, G., Taylor, M. C., Blundell, P. A., Van Leeuwen, F., Bitter, W., Cross, M. & McCulloch, R. (1996).** Antigenic variation in trypanosomes. *Archives of Medical Research* **27** (3), 379-388.
- Borst, P. & Fairlamb, A. H. (1998).** Surface receptors and transporters of *Trypanosoma brucei*. *Annual Review of Microbiology* **52** (1), 745-778.
- Borst, P. & Ulbert, S. (2001).** Control of VSG gene expression sites. *Molecular and Biochemical Parasitology* **114** (1), 17-27.
- Boulton, S. J. & Jackson, S. P. (1996).** Identification of a *Saccharomyces cerevisiae* Ku80 homologue: roles in DNA double strand break rejoining and in telomeric maintenance. *Nucleic Acids Research* **24** (23), 4639-4648.
- Boulton, S. J., Gartner, A., Reboul, J., Vaglio, P., Dyson, N., Hill, D. E. & Vidal, M. (2002).** Combined functional genomic maps of the *C. elegans* DNA damage response. *Science* **295** (5552), 127-131.
- Braybrooke, J. P., Spink, K. G., Thacker, J. & Hickson, I. D. (2000).** The RAD51 family member, RAD51L3, is a DNA-stimulated ATPase that forms a complex with XRCC2. *Journal of Biological Chemistry* **275** (37), 29100-29106.
- Braybrooke, J. P., Li, J.-L., Wu, L., Caple, F., Benson, F. E. & Hickson, I. D. (2003).** Functional Interaction between the Bloom's Syndrome Helicase and the RAD51 Paralog, RAD51L3 (RAD51D). *Journal of Biological Chemistry* **278** (48), 48357-48366.
- Breeden, L. & Nasmyth, K. (1985).** Regulation of the yeast HO gene. *Cold Spring Harbor symposia on quantitative biology* **50** 643-650.
- Brent, R. & Ptashne, M. (1980).** The *lexA* gene product represses its own promoter. *Proceedings of the National Academy of Sciences* **77** (4), 1932-1936.

- Brent, R. & Ptashne, M. (1985).** A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. *Cell* **43** 729-736.
- Burton, P., McBride, D. J., Wilkes, J. M., Barry, J. D. & McCulloch, R. (2007).** Ku heterodimer-independent end joining in *Trypanosoma brucei* cell extracts relies upon sequence microhomology. *Eukaryotic Cell* **6** (10), 1773-1781.
- Carlton, J. M., Hirt, R. P., Silva, J. C., Delcher, A. L., Schatz, M., Zhao, Q., Wortman, J. R., Bidwell, S. L., Alsmark, U. C. M., Besteiro, S., Sicheritz-Ponten, T., Noel, C. J., Dacks, J. B., Foster, P. G., Simillion, C., Van de Peer, Y., Miranda-Saavedra, D., Barton, G. J., Westrop, G. D., Muller, S., Dessi, D., Fiori, P. L., Ren, Q., Paulsen, I., Zhang, H., Bastida-Corcuera, F. D., Simoes-Barbosa, A., Brown, M. T., Hayes, R. D., Mukherjee, M., Okumura, C. Y., Schneider, R., Smith, A. J., Vanacova, S., Villalvazo, M., Haas, B. J., Pertea, M., Feldblyum, T. V., Utterback, T. R., Shu, C.-L., Osoegawa, K., de Jong, P. J., Hrdy, I., Horvathova, L., Zubacova, Z., Dolezal, P., Malik, S.-B., Logsdon, J. M., Jr., Henze, K., Gupta, A., Wang, C. C., Dunne, R. L., Upcroft, J. A., Upcroft, P., White, O., Salzberg, S. L., Tang, P., Chiu, C.-H., Lee, Y.-S., Embley, T. M., Coombs, G. H., Mottram, J. C., Tachezy, J., Fraser-Liggett, C. M. & Johnson, P. J. (2007).** Draft genome sequence of the sexually transmitted pathogen *Trichomonas vaginalis*. *Science* **315** (5809), 207-212.
- Carrington, M., Miller, N., Blum, M., Roditi, I., Wiley, D. & Turner, M. (1991).** Variant specific glycoprotein of *Trypanosoma brucei* consists of two domains each having an independently conserved pattern of cysteine residues. *Journal of Molecular Biology* **221** (3), 823-835.
- Cartwright, R., Dunn, A. M., Simpson, P. J., Tambini, C. E. & Thacker, J. (1998).** Isolation of novel human and mouse genes of the recA/RAD51 recombination-repair gene family. *Nucleic Acids Research* **26** (7), 1653-1659.
- Causier, B. & Davies, B. (2002).** Analysing protein-protein interactions with the yeast two-hybrid system. *Plant Molecular Biology* **50** (6), 855-870.
- Cerutti, H., Osman, M., Grandoni, P. & Jagendorf, A. T. (1992).** A homolog of *Escherichia coli* RecA protein in plastids of higher plants. *Proceedings of the National Academy of Sciences* **89** (17), 8068-8072.
- Chanet, R., Heude, M., Adjiri, A., Maloisel, L. & Fabre, F. (1996).** Semidominant mutations in the yeast Rad51 protein and their relationships with the Srs2 helicase. *Molecular and Cellular Biology* **16** (9), 4782-4789.
- Chang, D. J. & Cimprich, K. A. (2009).** DNA damage tolerance: when it's OK to make mistakes. *Nature Chemical Biology* **5** (2), 82-90.
- Chattopadhyay, A., Jones, N. G., Nietlispach, D., Nielsen, P. R., Voorheis, H. P., Mott, H. R. & Carrington, M. (2005).** Structure of the C-terminal domain from *Trypanosoma brucei* Variant Surface Glycoprotein MITat1.2. *Journal of Biological Chemistry* **280** (8), 7228-7235.
- Chaves, I., Rudenko, G., Dirks-Mulder, A., Cross, M. & Borst, P. (1999).** Control of variant surface glycoprotein gene-expression sites in *Trypanosoma brucei*. *The EMBO Journal* **18** (17), 4846-4855.

- Chen, Z., Yang, H. & Pavletich, N. P. (2008).** Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures. *Nature* **453** (7194), 489-494.
- Christmann, M., Tomicic, M. T., Roos, W. P. & Kaina, B. (2003).** Mechanisms of human DNA repair: an update. *Toxicology* **193** (1-2), 3-34.
- Clark, A. J. & Margulies, A. D. (1965).** Isolation and characterization of Recombination-Deficient Mutants of *Escherichia coli* K12. *Proceedings of the National Academy of Sciences* **53** 451-459.
- Claussen, C. A. & Long, E. C. (1999).** Nucleic acid recognition by metal complexes of bleomycin. *Chemical Reviews* **99** (9), 2797-2816.
- Clayton, C. E. (2002).** Life without transcriptional control? From fly to man and back again. *The EMBO Journal* **21** (8), 1881-1888.
- Collins, A. R. (1993).** Mutant rodent cell lines sensitive to ultraviolet light, ionizing radiation and cross-linking agents: a comprehensive survey of genetic and biochemical characteristics. *Mutation Research/DNA Repair* **293** 99-118.
- Conway, A. B., Lynch, T. W., Zhang, Y., Fortin, G. S., Fung, C. W., Symington, L. S. & Rice, P. A. (2004).** Crystal structure of a Rad51 filament. *Nature Structural & Molecular Biology* **11** (8), 791-796.
- Conway, C., McCulloch, R., Ginger, M. L., Robinson, N. P., Browitt, A. & Barry, J. D. (2002a).** Ku is important for telomere maintenance, but not for differential expression of telomeric VSG genes, in African Trypanosomes. *Journal of Biological Chemistry* **277** (24), 21269-21277.
- Conway, C., Proudfoot, C., Burton, P., Barry, J. D. & McCulloch, R. (2002b).** Two pathways of homologous recombination in *Trypanosoma brucei*. *Molecular Microbiology* **45** (6), 1687-1700.
- Corneo, B., Wendland, R. L., Deriano, L., Cui, X., Klein, I. A., Wong, S.-Y., Arnal, S., Holub, A. J., Weller, G. R., Pancake, B. A., Shah, S., Brandt, V. L., Meek, K. & Roth, D. B. (2007).** Rag mutations reveal robust alternative end joining. *Nature* **449** (7161), 483-486.
- Cox, M. M. (2000).** Recombinational DNA repair in bacteria and the RecA protein. *Progress in Nucleic Acid Research and Molecular Biology* **63** 311-366.
- Craig, N. L. & Roberts, J. W. (1981).** Function of nucleoside triphosphate and polynucleotide in *Escherichia coli* recA protein-directed cleavage of phage lambda repressor. *Journal of Biological Chemistry* **256** (15), 8039-8044.
- Critchlow, S. E., Bowater, R. P. & Jackson, S. P. (1997).** Mammalian DNA double-strand break repair protein XRCC4 interacts with DNA ligase IV. *Current Biology* **7** (8), 588-598.
- Cross, G. A. (1975).** Identification, purification and properties of clone-specific glycoprotein antigens constituting the surface coat of *Trypanosoma brucei*. *Parasitology* **71** 393-417.

**Cross, G. A. M. (2001).** African trypanosomes in the 21st century: what is their future in science and in health? *International Journal for Parasitology* **31** (5-6), 427-433.

**Dai, Q., Restrepo, B. I., Porcella, S. F., Raffel, S. J., Schwan, T. G. & Barbour, A. G. (2006).** Antigenic variation by *Borrelia hermsii* occurs through recombination between extragenic repetitive elements on linear plasmids. *Molecular Microbiology* **60** (6), 1329-1343.

**Dai, Y., Kysela, B., Hanakahi, L. A., Manolis, K., Riballo, E., Stumm, M., Harville, T. O., West, S. C., Oettinger, M. A. & Jeggo, P. A. (2003).** Nonhomologous end joining and V(D)J recombination require an additional factor. *Proceedings of the National Academy of Sciences* **100** (5), 2462-2467.

**Davis, A. P. & Symington, L. S. (2004).** RAD51-dependent Break-Induced Replication in yeast. *Molecular and Cellular Biology* **24** (6), 2344-2351.

**De Lange, T. & Borst, P. (1982).** Genomic environment of the expression-linked extra copies of genes for surface antigens of *Trypanosoma brucei* resembles the end of a chromosome. *Nature* **299** (5882), 451-453.

**De Lange, T., Kooter, J. M., Michels, P. A. M. & Borst, P. (1983).** Telomere conversion in trypanosomes. *Nucleic Acids Research* **11** (23), 8149-8165.

**Deans, B., Griffin, C. S., Maconochie, M. & Thacker, J. (2000).** Xrcc2 is required for genetic stability, embryonic neurogenesis and viability in mice. *The EMBO Journal* **19** (24), 6675-6685.

**Deitsch, K. W., Moxon, E. R. & Wellems, T. E. (1997).** Shared themes of antigenic variation and virulence in bacterial, protozoal, and fungal infections. *Microbiological & Molecular Biology Reviews* **61** (3), 281-293.

**Di Capua, E., Engel, A., Stasiak, A. & Koller, T. (1982).** Characterization of complexes between recA protein and duplex DNA by electron microscopy. *Journal of Molecular Biology* **157** (1), 87-103.

**Donelson, J. E. (2003).** Antigenic variation and the African trypanosome genome. *Acta Tropica* **85** (3), 391-404.

**Dosanji, M. K., Collins, D. W., Fan, W., Lennon, G. G., Albala, J. S., Shen, Z. & Schild, D. (1998).** Isolation and characterization of RAD51C, a new human member of the RAD51 family of related genes. *Nucleic Acids Research* **26** (5), 1179-1184.

**Doutriaux, M. P., Couteau, F., Bergounioux, C. & White, C. (1998).** Isolation and characterisation of the RAD51 and DMC1 homologs from *Arabidopsis thaliana*. *Molecular Genes and Genetics* **257** 283-291.

**Doyle, J. J., Hirumi, H., Hirumi, K., Lupton, E. N. & Cross, G. A. (1980).** Antigenic variation in clones of animal-infective *Trypanosoma brucei* derived and maintained *in vitro*. *Parasitology* **80** (2), 359-369.

**Dreesen, O., Li, B. & Cross, G. A. M. (2007).** Telomere structure and function in trypanosomes: a proposal. *Nature Reviews Microbiology* **5** (1), 70-75.

**Dudasova, Z., Dudas, A. & Chovanec, M. (2004).** Non-homologous end-joining factors of *Saccharomyces cerevisiae*. *FEMS Microbiology Reviews* **28** 581-601.

**Dudley, D. D., Chaudhuri, J., Bassing, C. H., Alt, F. W. & Frederick, W. A. (2005).** Mechanism and control of V(D)J recombination versus class switch recombination: similarities and differences. In *Advances in Immunology*, pp. 43-112: Academic Press.

**Dumas, P., Bergdoll, M., Cagnon, C. & Masson, J. M. (1994).** Crystal structure and site-directed mutagenesis of a bleomycin resistance protein and their significance for drug sequestering. *The EMBO Journal* **13** (11), 2483-2492.

**Eichinger, L., Pachebat, J. A., Glockner, G., Rajandream, M. A., Sucgang, R., Berriman, M., Song, J., Olsen, R., Szafranski, K., Xu, Q., Tunggal, B., Kummerfeld, S., Madera, M., Konfortov, B. A., Rivero, F., Bankier, A. T., Lehmann, R., Hamlin, N., Davies, R., Gaudet, P., Fey, P., Pilcher, K., Chen, G., Saunders, D., Sodergren, E., Davis, P., Kerhornou, A., Nie, X., Hall, N., Anjard, C., Hemphill, L., Bason, N., Farbrother, P., Desany, B., Just, E., Morio, T., Rost, R., Churcher, C., Cooper, J., Haydock, S., van Driessche, N., Cronin, A., Goodhead, I., Muzny, D., Mourier, T., Pain, A., Lu, M., Harper, D., Lindsay, R., Hauser, H., James, K., Quiles, M., Madan Babu, M., Saito, T., Buchrieser, C., Wardroper, A., Felder, M., Thangavelu, M., Johnson, D., Knights, A., Loulseged, H., Mungall, K., Oliver, K., Price, C., Quail, M. A., Urushihara, H., Hernandez, J., Rabbinoiwitsch, E., Steffen, D., Sanders, M., Ma, J., Kohara, Y., Sharp, S., Simmonds, M., Spiegler, S., Tivey, A., Sugano, S., White, B., Walker, D., Woodward, J., Winckler, T., Tanaka, Y., Shaulsky, G., Schleicher, M., Weinstock, G., Rosenthal, A., Cox, E. C., Chisholm, R. L., Gibbs, R., Loomis, W. F., Platzer, M., Kay, R. R., Williams, J., Dear, P. H., Noegel, A. A., Barrell, B. & Kuspa, A. (2005).** The genome of the social amoeba *Dictyostelium discoideum*. *Nature* **435** (7038), 43-57.

**Eisen, J. A. & Hanawalt, P. C. (1999).** A phylogenomic study of DNA repair genes, proteins, and processes. *Mutation Research/DNA Repair* **435** (3), 171-213.

**Eisen, J. A., Coyne, R. S., Wu, M., Wu, D., Thiagarajan, M., Wortman, J. R., Badger, J. H., Ren, Q., Amedeo, P., Jones, K. M., Tallon, L. J., Delcher, A. L., Salzberg, S. L., Silva, J. C., Haas, B. J., Majoros, W. H., Farzad, M., Carlton, J. M., Smith, R. K., Garg, J., Pearlman, R. E., Karrer, K. M., Sun, L., Manning, G., Elde, N. C., Turkewitz, A. P., Asai, D. J., Wilkes, D. E., Wang, Y., Cai, H., Collins, K., Stewart, B. A., Lee, S. R., Wilamowska, K., Weinberg, Z., Ruzzo, W. L., Wloga, D., Gaertig, J., Frankel, J., Tsao, C.-C., Gorovsky, M. A., Keeling, P. J., Waller, R. F., Patron, N. J., Cherry, J. M., Stover, N. A., Krieger, C. J., del Toro, C., Ryder, H. F., Williamson, S. C., Barbeau, R. A., Hamilton, E. P. & Orias, E. (2006).** Macronuclear genome sequence of the ciliate *Tetrahymena thermophila*, a model eukaryote. *Public Library of Science Biology* **4** (9), e286.

**El-Sayed, N. M., Myler, P. J., Bartholomeu, D. C., Nilsson, D., Aggarwal, G., Tran, A.-N., Ghedin, E., Worthey, E. A., Delcher, A. L., Blandin, G., Westenberger, S. J., Caler, E., Cerqueira, G. C., Branche, C., Haas, B., Anupama, A., Arner, E., Aslund, L., Attipoe, P., Bontempi, E., Brinkaud, F., Burton, P., Cadag, E., Campbell, D. A., Carrington, M., Crabtree, J., Darban, H., da Silveira, J. F., de Jong, P., Edwards, K., Englund, P. T., Fazelina, G., Feldblyum, T., Ferella, M., Frasch, A. C., Gull, K., Horn, D., Hou, L., Huang, Y., Kindlund, E., Klingbeil, M., Kluge, S., Koo, H., Lacerda, D., Levin, M. J., Lorenzi, H., Louie, T., Machado, C. R., McCulloch, R., McKenna, A., Mizuno, Y., Mottram, J. C., Nelson, S., Ochaya, S., Osoegawa, K., Pai, G., Parsons, M., Pentony, M., Pettersson, U., Pop, M., Ramirez, J. L., Rinta, J.,**



**Robertson, L., Salzberg, S. L., Sanchez, D. O., Seyler, A., Sharma, R., Shetty, J., Simpson, A. J., Sisk, E., Tammi, M. T., Tarleton, R., Teixeira, S., Van Aken, S., Vogt, C., Ward, P. N., Wickstead, B., Wortman, J., White, O., Fraser, C. M., Stuart, K. D. & Andersson, B. (2005).** The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas Disease. *Science* **309** (5733), 409-415.

**Enanga, B., Burchmore, R. J. S., Stewart, M. L. & Barrett, M. P. (2002).** Sleeping sickness and the brain. *Cellular and Molecular Life Sciences* **59** (5), 845-858.

**Engstler, M., Pfohl, T., Herminghaus, S., Boshart, M., Wiegertjes, G., Heddergott, N. & Overath, P. (2007).** Hydrodynamic flow-mediated protein sorting on the cell surface of trypanosomes. *Cell* **131** (3), 505-515.

**Evans, D. & Blumenthal, T. (2000).** *trans*-Splicing of polycistronic *Caenorhabditis elegans* pre-mRNAs: Analysis of the SL2 RNA. *Molecular and Cellular Biology* **20** (18), 6659-6667.

**Falck, J., Coates, J. & Jackson, S. P. (2005).** Conserved modes of recruitment of ATM, ATR and DNA-PKcs to sites of DNA damage. *Nature* **434** (7033), 605-611.

**Falzon, M., Fewell, J. W. & Kuff, E. L. (1993).** EBP-80, a transcription factor closely resembling the human autoantigen Ku, recognizes single- to double-strand transitions in DNA. *Journal of Biological Chemistry* **268** (14), 10546-10552.

**Fasullo, M., Dave, P. & Rothstein, R. (1994).** DNA-damaging agents stimulate the formation of directed reciprocal translocations in *Saccharomyces cerevisiae*. *Mutation Research* **314** (2), 121-133.

**Fenn, K. & Matthews, K. R. (2007).** The cell biology of *Trypanosoma brucei* differentiation. *Current Opinion in Microbiology* **10** (6), 539-546.

**Ferguson, D. O. & Holloman, W. K. (1996).** Recombinational repair of gaps in DNA is asymmetric in *Ustilago maydis* and can be explained by a migrating D-loop model. *Proceedings of the National Academy of Sciences* **93** (11), 5419-5424.

**Ferguson, D. O., Rice, M. C., Rendi, M. H., Kotani, H., Kmiec, E. B. & Holloman, W. K. (1997).** Interaction between *Ustilago maydis* REC2 and RAD51 genes in DNA repair and mitotic recombination. *Genetics* **145** (2), 243-251.

**Ferguson, M. A., Homans, S. W., Dwek, R. A. & Rademacher, T. W. (1988).** Glycosyl-phosphatidylinositol moiety that anchors *Trypanosoma brucei* variant surface glycoprotein to the membrane. *Science* **239** (4841), 753-759.

**Field, M. C. & Boothroyd, J. C. (1996).** Sequence divergence in a family of variant surface glycoprotein genes from trypanosomes: coding region hypervariability and downstream recombinogenic repeats. *Journal of Molecular Evolution* **42** (5), 500-511.

**Fields, S. & Song, O. (1989).** A novel genetic system to detect protein – protein interactions. *Nature* **340** (6230), 245-246.

**Figueiredo, L. M., Janzen, C. J. & Cross, G. A. M. (2008).** A histone methyltransferase modulates antigenic variation in African Trypanosomes. *Public Library of Science, Biology* **6** (7), e161.

- Fishman-Lobell, J., Rudin, N. & Haber, J. E. (1992).** Two alternative pathways of double-strand break repair that are kinetically separable and independently modulated. *Molecular and Cellular Biology* **12** (3), 1292-1303.
- Friedberg, E. C., Walker, G. C. & Siede, W. (1995).** *DNA repair and Mutagenesis*: ASM Press, Washington D.C.
- Fritsch, O., Benvenuto, G., Bowler, C., Molinier, J. & Hohn, B. (2004).** The INO80 protein controls homologous recombination in *Arabidopsis thaliana*. *Molecular Cell* **16** (3), 479-485.
- Gajria, B., Bahl, A., Brestelli, J., Dommer, J., Fischer, S., Gao, X., Heiges, M., Iodice, J., Kissinger, J. C., Mackey, A. J., Pinney, D. F., Roos, D. S., Stoeckert, C. J., Jr., Wang, H. & Brunk, B. P. (2008).** ToxoDB: an integrated *Toxoplasma gondii* database resource. *Nucleic Acids Research* **36** 553-556.
- Galletto, R., Amitani, I., Baskin, R. J. & Kowalczykowski, S. C. (2006).** Direct observation of individual RecA filaments assembling on single DNA molecules. *Nature* **443** (7113), 875-878.
- Game, J. C. & Mortimer, R. K. (1974).** Genetic study of x-ray sensitive mutants in yeast. *Mutation Research* **24** (3), 281-292.
- Gasior, S. L., Wong, A. K., Kora, Y., Shinohara, A. & Bishop, D. K. (1998).** Rad52 associates with RPA and functions with Rad55 and Rad57 to assemble meiotic recombination complexes. *Genes & Development* **12** 2208-2221.
- Gasior, S. L., Olivares, H., Ear, U., Hari, D. M., Weichselbaum, R. & Bishop, D. K. (2001).** Assembly of RecA-like recombinases: distinct roles for mediator proteins in mitosis and meiosis. *Proceedings of the National Academy of Sciences* **98** (15).
- Gellert, M. (2002).** V(D)J recombination: RAG proteins, repair factor and regulation. *Annual Review of Biochemistry* **71** (1), 101-132.
- geneDB (2009).** [www.genedb.org/](http://www.genedb.org/).
- Ghabrial, A., Ray, R. P. & Schupbach, T. (1998).** okra and spindle-B encode components of the RAD52 DNA repair pathway and affect meiosis and patterning in *Drosophila* oogenesis. *Genes & Development* **12** (17), 2711-2723.
- Gibson, W., Peacock, L., Ferris, V., Williams, K. & Bailey, M. (2008).** The use of yellow fluorescent hybrids to indicate mating in *Trypanosoma brucei*. *Parasites & Vectors* **1** (1), 4.
- Gibson, W. C. (1995).** The significance of genetic exchange in trypanosomes. *Parasitology Today* **11** (12), 465-468.
- Gill, E. E. & Fast, N. M. (2007).** Stripped-down DNA repair in a highly reduced parasite. *BMC Molecular Biology* **8** 24.
- Giloni, L., Takeshita, M., Johnson, F., Iden, C. & Grollman, A. P. (1981).** Bleomycin-induced strand-scission of DNA: Mechanism of deoxyribose cleavage. *Journal of Biological Chemistry* **256** (16), 8608-8615.

**Giniger, E., Varnum, S. M. & Ptashne, M. (1985).** Specific DNA binding of GAL4, a positive regulatory protein of yeast. *Cell* **40** (4), 767-774.

**Glover, L., Alsford, S., Beattie, C. & Horn, D. (2007).** Deletion of a trypanosome telomere leads to loss of silencing and progressive loss of terminal DNA in the absence of cell cycle arrest. *Nucleic Acids Research* **35** (3), 872-880.

**Glover, L., McCulloch, R. & Horn, D. (2008).** Sequence homology and microhomology dominate chromosomal double-strand break repair in African trypanosomes. *Nucleic Acids Research* **36** (8), 2608-2618.

**Goodarzi, A. A., Yu, Y., Riballo, E., Douglas, P., Walker, S. A., Ye, R., Harer, C., Marchetti, C., Morrice, N., Jeggo, P. A. & Lees-Miller, S. P. (2006).** DNA-PK autophosphorylation facilitates Artemis endonuclease activity. *The EMBO Journal* **25** (16), 3880-3889.

**Gottlieb, T. M. & Jackson, S. P. (1993).** The DNA-dependent protein kinase: Requirement for DNA ends and association with Ku antigen. *Cell* **72** (1), 131-142.

**Grishchuk, A. L. & Kohli, J. (2003).** Five RecA-like Proteins of *Schizosaccharomyces pombe* are involved in Meiotic Recombination. *Genetics* **165** (3), 1031-1043.

**Gull, K. (2002).** The cell biology of parasitism in *Trypanosoma brucei*: insights and drug targets from genomic approaches? *Current Pharmaceutical Design* **8** (4), 241-256.

**Guy, C. P., Haldenby, S., Brindley, A., Walsh, D. A., Briggs, G. S., Warren, M. J., Allers, T. & Bolt, E. L. (2006).** Interactions of RadB, a DNA repair protein in archaea, with DNA and ATP. *Journal of Molecular Biology* **358** (1), 46-56.

**Gyuris, J., Golemis, E., Chertkov, H. & Brent, R. (1993).** Cdi1, a human G1 and S phase protein phosphatase that associates with Cdk2. *Cell* **75** (4), 791-803.

**Haaf, T., Golub, E. I., Reddy, G., Radding, C. M. & Ward, D. C. (1995).** Nuclear foci of mammalian Rad51 recombination protein in somatic cells after DNA damage and its localization in synaptonemal complexes. *Proceedings of the National Academy of Sciences* **92** (6), 2298-2302.

**Haaf, T., Raderschall, E., Reddy, G., Ward, D. C., Radding, C. M. & Golub, E. I. (1999).** Sequestration of mammalian Rad51-recombination protein into micronuclei. *Journal of Cell Biology* **144** (1), 11-20.

**Haber, J., Ira, G., Malkova, A. & Sugawara, N. (2004).** Repairing a double-strand chromosome break by homologous recombination: revisiting Robin Holliday's model. *Philosophical Transactions of the Royal Society of London* **359** (1441), 79-86.

**Haber, J. E. (1992).** Mating-type gene switching in *Saccharomyces cerevisiae*. *Trends in Genetics* **8** (12), 446-452.

**Haber, J. E. (2006).** Chromosome Breakage and Repair. *Genetics* **173** (3), 1181-1185.

**Hakem, R. (2008).** DNA-damage repair; the good, the bad, and the ugly. *The EMBO Journal* **27** (4), 589-605.

- Hartley, C. L. (2008).** Characterisation of factors that regulate homologous recombination and antigenic variation in *Trypanosoma brucei*, pp. 392: *PhD Thesis In: Wellcome Centre for Molecular Parasitology*, University of Glasgow, Glasgow, UK.
- Hartley, C. L. & McCulloch, R. (2008).** *Trypanosoma brucei* BRCA2 acts in antigenic variation and has undergone a recent expansion in BRC repeat number that is important during homologous recombination. *Molecular Microbiology* **68** (5), 1237-1251.
- Hasegawa, Y., Wakabayashi, M., Nakamura, S., Kodaira, K.-i., Shinohara, H. & Yasukawa, H. (2004).** A homolog of *Escherichia coli* RecA in mitochondria of the cellular slime mold *Dictyostelium discoideum*. *DNA Repair* **3** (5), 515-525.
- Hays, S. L., Firmenich, A. A. & Berg, P. (1995).** Complex formation in yeast double-strand break repair: Participation of Rad51, Rad52, Rad55, and Rad57 proteins. *Proceedings of the National Academy of Sciences* **92** (15), 6925-6929.
- Heacock, M., Spangler, E., Riha, K., Puizina, J. & Shippen, D. E. (2004).** Molecular analysis of telomere fusions in *Arabidopsis*: multiple pathways for chromosome end-joining. *The EMBO Journal* **23** (11), 2304-2313.
- Hedges, S., Blair, J., Venturi, M. & Shoe, J. (2004).** A molecular timescale of eukaryote evolution and the rise of complex multicellular life. *BMC Evolutionary Biology* **4** (1), 2.
- Hedges, S. B. (2002).** The origin and evolution of model organisms. *Nature Reviews Genetics* **3** (11), 838-849.
- Hefferin, M. L. & Tomkinson, A. E. (2005).** Mechanism of DNA double-strand break repair by non-homologous end joining. *DNA Repair* **4** 639-648.
- Hertz-Fowler, C., Renauld, H. & Berriman, M. (2007).** The genome of *Trypanosoma brucei*. In *Trypanosomes: After the Genome*, pp. 5-48. Edited by D. Barry, R. McCulloch, J. C. Mottram & A. Acosta-Serrano. Norfolk, UK: Horizon Bioscience.
- Hertz-Fowler, C., Figueiredo, L. M., Quail, M. A., Becker, M., Jackson, A., Bason, N., Brooks, K., Churcher, C., Fahkro, S., Goodhead, I., Heath, P., Kartvelishvili, M., Mungall, K., Harris, D., Hauser, H., Sanders, M., Saunders, D., Seeger, K., Sharp, S., Taylor, J. E., Walker, D., White, B., Young, R., Cross, G. A. M., Rudenko, G., Barry, J. D., Louis, E. J. & Berriman, M. (2008).** Telomeric expression sites are highly conserved in *Trypanosoma brucei*. *Public Library of Science, ONE* **3** (10), 3527.
- Hide, G., Cattand, P., LeRay, D., Barry, J. D. & Tait, A. (1990).** The identification of *Trypanosoma brucei* subspecies using repetitive DNA sequences. *Molecular and Biochemical Parasitology* **39** (2), 213-225.
- Hide, G. (1999).** History of Sleeping Sickness in East Africa. *Clinical Microbiology Reviews* **12** (1), 112-125.
- Hirumi, H. & Hirumi, K. (1994).** Axenic culture of African trypanosome bloodstream forms. *Parasitology Today* **10** (2), 80-84.
- Hoeijmakers, J. H., Frasch, A. C., Bernardis, A., Borst, P. & Cross, G. A. (1980).** Novel expression-linked copies of the genes for variant surface antigens in trypanosomes. *Nature* **284** (5751), 78-80.

**Holloman, W. K., Schirawski, J. & Holliday, R. (2008).** The homologous recombination system of *Ustilago maydis*. *Fungal Genetics and Biology* **45** S31-S39.

**Hunter, P. (2008).** The paradox of model organisms. *EMBO reports* **9** (8), 717-720.

**Ip, S. C. Y., Rass, U., Blanco, M. G., Flynn, H. R., Skehel, J. M. & West, S. C. (2008).** Identification of Holliday junction resolvases from humans and yeast. *Nature* **456** (7220), 357-361.

**Ira, G. & Haber, J. E. (2002).** Characterization of RAD51-independent break-induced replication that acts preferentially with short homologous sequences. *Molecular and Cellular Biology* **22** (18), 6384-6392.

**Ishibashi, T., Isogai, M., Kiyohara, H., Hosaka, M., Chiku, H., Koga, A., Yamamoto, T., Uchiyama, Y., Mori, Y., Hashimoto, J., Ausio, J., Kimura, S. & Sakaguchi, K. (2006).** Higher plant RecA-like protein is homologous to RadA. *DNA Repair* **5** (1), 80-88.

**Ivanov, E. L., Sugawara, N., Fishman-Lobell, J. & Haber, J. E. (1996).** Genetic requirements for the single-strand annealing pathway of Double-Strand Break repair in *Saccharomyces cerevisiae*. *Genetics* **142** (3), 693-704.

**Jain, S. K., Cox, M. M. & Inman, R. B. (1994).** On the role of ATP hydrolysis in RecA protein-mediated DNA strand exchange. III. Unidirectional branch migration and extensive hybrid DNA formation. *Journal of Biological Chemistry* **269** (32), 20653-20661.

**Jannin, J. & Cattand, P. (2004).** Treatment and control of human African trypanosomiasis. *Current Opinion in Infectious Disease* **17** (6), 565-571.

**Johnson, P. J., Kooter, J. M. & Borst, P. (1987).** Inactivation of transcription by UV irradiation of *T. brucei* provides evidence for a multicistronic transcription unit including a VSG gene. *Cell* **51** (2), 273-281.

**Johnson, R. D. & Symington, L. S. (1995).** Functional differences and interactions among the putative RecA homologs Rad51, Rad55, and Rad57. *Molecular and Cellular Biology* **15** (9), 4843-4850.

**Johnson, R. D., Liu, N. & Jasin, M. (1999).** Mammalian XRCC2 promotes the repair of DNA double-strand breaks by homologous recombination. *Nature* **401** 397-399.

**Jones, N. J., Cox, R. & Thacker, J. (1987).** Isolation and cross-sensitivity of X-ray-sensitive mutants of V79-4 hamster cells. *Mutation Research* **183** (3), 279-286.

**Jones, N. J., Cox, R. & Thacker, J. (1988).** Six complementation groups for ionising-radiation sensitivity in chinese hamster cells. *Mutation Research* **193** (2), 139-144.

**Kamper, S. M. & Barbet, A. F. (1992).** Surface epitope variation via mosaic gene formation is potential key to long-term survival of *Trypanosoma brucei*. *Molecular and Biochemical Parasitology* **53** 33-44.

**Kanazawa, S., Driscoll, M. & Struhl, K. (1988).** ATR1, a *Saccharomyces cerevisiae* gene encoding a transmembrane protein required for aminotriazole resistance. *Molecular and Cellular Biology* **8** (2), 664-673.

**Katinka, M. D., Duprat, S., Cornillot, E., Metenier, G., Thomarat, F., Prensier, G., Barbe, V., Peyretailade, E., Brottier, P., Wincker, P., Delbac, F., El Alaoui, H., Peyret, P., Saurin, W., Gouy, M., Weissenbach, J. & Vivares, C. P. (2001).** Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*. *Nature* **414** (6862), 450-453.

**Kawabata, M., Kawabata, T. & Nishibori, M. (2005).** Role of recA/RAD51 family proteins in mammals. *Acta Medica Okayama* **59** (1), 1-9.

**Keeney, S., Giroux, C. N. & Kleckner, N. (1997).** Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. *Cell* **88** (3), 375-384.

**Kelly, S., Reed, J., Kramer, S., Ellis, L., Webb, H., Sunter, J., Salje, J., Marinsek, N., Gull, K., Wickstead, B. & Carrington, M. (2007).** Functional genomics in *Trypanosoma brucei*: A collection of vectors for the expression of tagged proteins from endogenous and ectopic gene loci. *Molecular and Biochemical Parasitology* **154** (1), 103-109.

**Kennedy, P. G. (2004).** Human African trypanosomiasis of the CNS: current issues and challenges. *The Journal of Clinical Investigation* **113** (4), 496-5051.

**Khasanov, F. K., Salakhova, A. F., Chepurnaja, O. V., Korolev, V. G. & Bashkirov, V. I. (2004).** Identification and characterization of the rlp1+, the novel Rad51 paralog in the fission yeast *Schizosaccharomyces pombe*. *DNA Repair* **3** (10), 1363-1374.

**Klar, A. J. (1987).** The mother-daughter mating type switching asymmetry of budding yeast is not conferred by the segregation of parental HO gene DNA strands. *Genes & Development* **1** (10), 1059-1064.

**Klingbeil, M. M., Burton, P., Barnes, R. & McCulloch, R. (2007).** The three Rs of the Trypanosomatid genomes: replication, recombination, and repair. In *Trypanosomes: After the Genome*, pp. 133-176. Edited by D. Barry, R. McCulloch, J. C. Mottram & A. Acosta-Serrano. Norfolk, UK: Horizon Bioscience.

**Kmiec, E. B., Cole, A. & Holloman, W. K. (1994).** The REC2 gene encodes the homologous pairing protein of *Ustilago maydis*. *Molecular and Cellular Biology* **14** (11), 7163-7172.

**Kraus, E., Leung, W.-Y. & Haber, J. E. (2001).** Break-induced replication: A review and an example in budding yeast. *Proceedings of the National Academy of Sciences* **98** (15), 8255-8262.

**Kunz, C., Saito, Y. & Schär, P. (2009).** Mismatched repair: variations on a theme. *Cellular and Molecular Life Sciences*.

**Kupiec, M. & Petes, T. D. (1988).** Meiotic recombination between repeated transposable elements in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology* **8** (7), 2942-2954.

**Kurimasa, A., Kumano, S., Boubnov, N. V., Story, M. D., Tung, C.-S., Peterson, S. R. & Chen, D. J. (1999).** Requirement for the kinase activity of human DNA-dependent protein kinase catalytic subunit in DNA strand break rejoining. *Molecular and Cellular Biology* **19** (5), 3877-3884.

- Kurumizaka, H., Ikawa, S., Nakada, M., Eda, K., Kagawa, W., Takata, M., Takeda, S., Yokoyama, S. & Shibata, T. (2001).** Homologous-pairing activity of the human DNA-repair proteins Xrcc3-Rad51C. *Proceedings of the National Academy of Sciences* **98** (10), 5538-5543.
- Kuznetsov, S., Pellegrini, M., Shuda, K., Fernandez-Capetillo, O., Liu, Y., Martin, B. K., Burkett, S., Southon, E., Pati, D., Tessarollo, L., West, S. C., Donovan, P. J., Nussenzweig, A. & Sharan, S. K. (2007).** RAD51C deficiency in mice results in early prophase I arrest in males and sister chromatid separation at metaphase II in females. *Journal of Cellular Biology* **176** (5), 581-592.
- Lao, J. P., Oh, S. D., Shinohara, M., Shinohara, A. & Hunter, N. (2008).** Rad52 promotes postinvasion steps of meiotic Double-Strand-Break repair. *Molecular Cell* **29** (4), 517-524.
- Lee, K. & Lee, S. E. (2007).** *Saccharomyces cerevisiae* Sae2- and Tel1-dependent single-strand DNA formation at DNA break promotes Microhomology-Mediated End Joining. *Genetics* **176** (4), 2003-2014.
- Lehninger, A. L., Nelson, D. L. & Cox, M. M. (1993).** *Principles of Biochemistry*, Second edn: Worth Publisher, USA.
- Li, X. & Heyer, W. (2008).** Homologous recombination in DNA repair and DNA damage tolerance. *Cell Research* **18** (1), 99-113.
- Li, Z., Golub, E. I., Gupta, R. & Radding, C. M. (1997).** Recombination activities of HsDmcl protein, the meiotic human homolog of RecA protein. *Proceedings of the National Academy of Sciences* **94** (21), 11221-11226.
- Liang, F., Romanienko, P. J., Weaver, D. T., Jeggo, P. A. & Jasin, M. (1996).** Chromosomal double-strand break repair in Ku80-deficient cells. *Proceedings of the National Academy of Sciences* **93** (17), 8929-8933.
- Liang, X.-h., Haritan, A., Uliel, S. & Michaeli, S. (2003).** *trans* and *cis* splicing in Trypanosomatids: mechanism, factors, and regulation. *Eukaryotic Cell* **2** (5), 830-840.
- Lichten, M. (2001).** Meiotic recombination: Breaking the genome to save it. *Current Biology* **11** (7), 253-256.
- Lieber, M. R., Ma, Y., Pannicke, U. & Schwarz, K. (2004).** The mechanism of vertebrate nonhomologous DNA end joining and its role in V(D)J recombination. *DNA Repair* **3** (8), 817-826.
- Lim, D. S. & Hasty, P. (1996).** A mutation in mouse rad51 results in an early embryonic lethal that is suppressed by a mutation in p53. *Molecular and Cellular Biology* **16** (12), 7133-7143.
- Lin, F. L., Sperle, K. & Sternberg, N. (1990).** Intermolecular recombination between DNAs introduced into mouse L cells is mediated by a nonconservative pathway that leads to crossover products. *Molecular and Cellular Biology* **10** (1), 103-112.
- Lin, Z., Kong, H., Nei, M. & Ma, H. (2006).** Origins and evolution of the recA/RAD51 gene family: Evidence for ancient gene duplication and endosymbiotic gene transfer. *Proceedings of the National Academy of Sciences* **103** (27), 10328-10333.

**Lio, Y. C., Mazin, A. V., Kowalczykowski, S. C. & Chen, D. J. (2003).** Complex formation by the human Rad51B and Rad51C DNA repair proteins and their activities *in-vitro*. *Journal of Biological Chemistry* **278** (4), 2469-2478.

**Lio, Y. C., Schild, D., Brenneman, M. A., Redpath, J. L. & Chen, D. J. (2004).** Human Rad51C deficiency destabilizes XRCC3, impairs recombination, and radiosensitizes S/G2-phase cells. *Journal of Biological Chemistry* **279** (40), 42313-42320.

**Lisby, M., Barlow, J. H., Burgess, R. C. & Rothstein, R. (2004).** Choreography of the DNA damage response: Spatiotemporal relationships among checkpoint and repair proteins. *Cell* **118** (6), 699-713.

**Little, J. W. & Mount, D. W. (1982).** The SOS regulatory system of *Escherichia coli*. *Cell* **29** (1), 11-22.

**Liu, A. Y., Van der Ploeg, L. H., Rijsewijk, F. A. & Borst, P. (1983).** The transposition unit of variant surface glycoprotein gene 118 of *Trypanosoma brucei*. Presence of repeated elements at its border and absence of promoter-associated sequences. *Journal of Molecular Biology* **167** (1), 57-75.

**Liu, N., Lamerdin, J. E., Tebbs, R. S., Schild, D., Tucker, J. D., Shen, M. R., Brookman, K. W., Siciliano, M. J., Walter, C. A., Fan, W., Narayana, L. S., Zhou, Z.-Q., Adamson, A. W., Sorensen, K. J., Chen, D. J., Jones, N. J. & Thompson, L. H. (1998).** XRCC2 and XRCC3, new human Rad51-family members, promote chromosome stability and protect against DNA cross-links and other damages. *Molecular Cell* **1** (6), 783-793.

**Liu, N., Schild, D., Thelen, M. P. & Thompson, L. H. (2002a).** Involvement of Rad51C in two distinct protein complexes of Rad51 paralogs in human cells. *Nucl Acids Res* **30** (4), 1009-1015.

**Liu, N., Schild, D., Thelen, M. P. & Thompson, L. H. (2002b).** Involvement of Rad51C in two distinct protein complexes of Rad51 paralogs in human cells. *Nucleic Acids Research* **30** (4), 1009-1015.

**Liu, Y. & Maizels, N. (2000).** Coordinated response of mammalian Rad51 and Rad52 to DNA damage. *EMBO reports* **1** (1), 85-90.

**Liu, Y., Masson, J.-Y., Shah, R., O'Regan, P. & West, S. C. (2004).** RAD51C is required for Holliday Junction processing in mammalian cells. *Science* **303** (5655), 243-246.

**Liu, Y., Tarsounas, M., O'Regan, P. & West, S. C. (2007).** Role of RAD51C and XRCC3 in genetic recombination and DNA repair. *Journal of Biological Chemistry* **282** (3), 1973-1979.

**Lo, T., Pellegrini, L., Venkitaraman, A. R. & Blundell, T. L. (2003).** Sequence fingerprints in BRCA2 and RAD51: implications for DNA repair and cancer. *DNA Repair* **2** (9), 1015-1028.

**Loftus, B., Anderson, I., Davies, R., Alsmark, U. C., Samuelson, J., Amedeo, P., Roncaglia, P., Berriman, M., Hirt, R. P., Mann, B. J., Nozaki, T., Suh, B., Pop, M., Duchene, M., Ackers, J., Tannich, E., Leippe, M., Hofer, M., Bruchhaus, I., Willhoeft,**



U., Bhattacharya, A., Chillingworth, T., Churcher, C., Hance, Z., Harris, B., Harris, D., Jagels, K., Moule, S., Mungall, K., Ormond, D., Squares, R., Whitehead, S., Quail, M. A., Rabinowitsch, E., Norbertczak, H., Price, C., Wang, Z., Guillen, N., Gilchrist, C., Stroup, S. E., Bhattacharya, S., Lohia, A., Foster, P. G., Sicheritz-Ponten, T., Weber, C., Singh, U., Mukherjee, C., El-Sayed, N. M., Petri, W. A., Jr., Clark, C. G., Embley, T. M., Barrell, B., Fraser, C. M. & Hall, N. (2005). The genome of the protist parasite *Entamoeba histolytica*. *Nature* **433** (7028), 865-868.

Lovett, S. T. & Mortimer, R. K. (1987). Characterization of null mutants of the RAD55 gene of *Saccharomyces cerevisiae*: Effects of temperature, osmotic strength and mating type. *Genetics* **116** (4), 547-553.

Luban, J. & Goff, S. P. (1995). The yeast two-hybrid system for studying protein-protein interactions. *Current Opinion in Biotechnology* **6** (1), 59-64.

Lukes, J., Lys Guilbride, D., Votypka, J., Zikova, A., Benne, R. & Englund, P. T. (2002). Kinetoplast DNA network: Evolution of an improbable structure. *Eukaryotic Cell* **1** (4), 495-502.

Lundin, C., North, M., Erixon, K., Walters, K., Jenssen, D., Goldman, A. S. H. & Helleday, T. (2005). Methyl methanesulfonate (MMS) produces heat-labile DNA damage but no detectable in vivo DNA double-strand breaks. *Nucleic Acids Research* **33** (12), 3799-3811.

Lundkvist, G. B., Kristensson, K. & Bentivoglio, M. (2004). Why trypanosomes cause sleeping sickness. *Physiology* **19** (4), 198-206.

Ma, J.-L., Kim, E. M., Haber, J. E. & Lee, S. E. (2003). Yeast Mre11 and Rad1 proteins define a Ku-independent mechanism to repair Double-Strand Breaks lacking overlapping end sequences. *Molecular and Cellular Biology* **23** (23), 8820-8828.

Ma, J. & Ptashne, M. (1987a). Deletion analysis of GAL4 defines two transcriptional activating segments. *Cell* **48** (5), 847-853.

Ma, J. & Ptashne, M. (1987b). A new class of yeast transcriptional activators. *Cell* **51** (1), 113-119.

MacLeod, A., Tweedie, A., McLellan, S., Taylor, S., Hall, N., Berriman, M., El-Sayed, N. M., Hope, M., Turner, C. M. R. & Tait, A. (2005). The genetic map and comparative analysis with the physical map of *Trypanosoma brucei*. *Nucleic Acids Research* **33** (21), 6688-6693.

MacLeod, A., Turner, M. & Tait, A. (2007). The system of genetic exchange in *Trypanosoma brucei* and other Trypanosomatids. In *Trypanosomes: After the genome*, pp. 71-90. Edited by D. Barry, R. McCulloch, J. C. Mottram & A. Acosta-Serrano. Norfolk, UK: Horizon Bioscience.

Malkova, A., Signon, L., Schaefer, C. B., Naylor, M. L., Theis, J. F., Newlon, C. S. & Haber, J. E. (2001). RAD51-independent break-induced replication to repair a broken chromosome depends on a distant enhancer site. *Genes & Development* **15** (9), 1055-1060.

Maloisel, L., Bhargava, J. & Roeder, G. S. (2004). A role for DNA polymerase  $\delta$  in gene conversion and crossing over during meiosis in *Saccharomyces cerevisiae*. *Genetics* **167** (3), 1133-1142.

- Malone, R. E. & Esposito, R. E. (1980).** The *RAD52* gene is required for homothallic interconversion of mating types and spontaneous mitotic recombination in yeast. *Proceedings of the National Academy of Sciences* **77** (1), 503-507.
- Mange, E. J. & Mange, A. P. (1999).** Basic Human Genetics. pp. 530. Massachusetts, USA: Sinauer.
- Marcello, L. & Barry, J. D. (2007a).** Analysis of the *VSG* gene silent archive in *Trypanosoma brucei* reveals that mosaic gene expression is prominent in antigenic variation and is favored by archive substructure. *Genome Research* **17** (9), 1344-1352.
- Marcello, L. & Barry, J. D. (2007b).** From silent genes to noisy populations; dialogue between the genotype and phenotypes of antigenic variation. *The Journal of Eukaryotic Microbiology* **54** (1), 14-17.
- Marnett, L. J. & Plataras, J. P. (2001).** Endogenous DNA damage and mutation. *Trends in Genetics* **17** (4), 214-221.
- Martin, J. S., Winkelmann, N., Petalcorin, M. I. R., McIlwraith, M. J. & Boulton, S. J. (2005).** RAD51-dependent and -independent roles of a *Caenorhabditis elegans* BRCA2-related protein during DNA double-strand break repair. *Molecular and Cellular Biology* **25** (8), 3127-3139.
- Martin, V., Chahwan, C., Gao, H., Blais, V., Wohlschlegel, J., Yates, J. R., McGowan, C. H. & Russell, P. (2006).** Sws1 is a conserved regulator of homologous recombination in eukaryotic cells. *EMBO journal* **25** (11), 2564-2574.
- Maslov, D. A., Podlipaev, S. A. & Lukes, J. (2001).** Phylogeny of the kinetoplastida: taxonomic problems and insights into the evolution of parasitism. *Memórias do Instituto Oswaldo Cruz* **96** (3), 397-402.
- Masson, J. Y., Stasiak, A. Z., Stasiak, A., Benson, F. E. & West, S. C. (2001a).** Complex formation by the human RAD51C and XRCC3 recombination repair proteins. *Proceedings of the National Academy of Sciences* **98** (15), 8440-8446.
- Masson, J. Y., Tarsounas, M. C., Stasiak, A. Z., Stasiak, A., Shah, R., McIlwraith, M. J., Benson, F. E. & West, S. C. (2001b).** Identification and purification of two distinct complexes containing the five RAD51 paralogs. *Genes and Development* **15** (24), 3296-3307.
- Matthews, K. R. & Gull, K. (1994).** Cycles within cycles: The interplay between differentiation and cell division in *Trypanosoma brucei*. *Parasitology Today* **10** (12), 473-476.
- Matthews, K. R., Tschudi, C. & Ullu, E. (1994).** A common pyrimidine-rich motif governs *trans*-splicing and polyadenylation of tubulin polycistronic pre-mRNA in trypanosomes. *Genes and Development* **8** (4), 491-501.
- Matthews, K. R., Ellis, J. R. & Paterou, A. (2004).** Molecular regulation of the life cycle of African trypanosomes. *Trends in Parasitology* **20** (1), 40-47.

- McCulloch, R., Rudenko, G. & Borst, P. (1997).** Gene conversions mediating antigenic variation in *Trypanosoma brucei* can occur in variant surface glycoprotein expression sites lacking 70-base-pair repeat sequences. *Molecular and Cellular Biology* **17** (2), 833-843.
- McCulloch, R. & Barry, J. D. (1999).** A role for RAD51 and homologous recombination in *Trypanosoma brucei* antigenic variation. *Genes & Development* **13** (21), 2875-2888.
- McCulloch, R. (2004).** Antigenic variation in African trypanosomes: monitoring progress. *Trends in Parasitology* **20** (3), 117-121.
- McEachern, M. J. & Haber, J. E. (2006).** Break-induced replication and recombinational telomere elongation in yeast. *Annual Review Biochemistry* **75** 111-135.
- McGhee, R. B. & Cosgrove, W. B. (1980).** Biology and physiology of the lower Trypanosomatidae. *Microbiological & Molecular Biology Reviews* **44** (1), 140-173.
- McGlynn, P. & Lloyd, R. G. (2002).** Recombinational repair and restart of damaged replication forks. *Nature Reviews Molecular Cell Biology* **3** (11), 859-870.
- McKean, P. G., Keen, J. K., Smith, D. F. & Benson, F. E. (2001).** Identification and characterisation of a RAD51 gene from *Leishmania major*. *Molecular and Biochemical Parasitology* **115** (2), 209-216.
- McKean, P. G. (2003).** Co-ordination of cell cycle and cytokinesis in *Trypanosoma brucei*. *Current Opinion in Microbiology* **6** (6), 600-607.
- McVey, M. & Lee, S. E. (2008).** MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. *Trends in Genetics* **24** (11), 529-538.
- Melville, S. E., Leech, V., Navarro, M. & Cross, G. A. M. (2000).** The molecular karyotype of the megabase chromosomes of *Trypanosoma brucei* stock 427. *Molecular and Biochemical Parasitology* **111** (2), 261-273.
- Michiels, F., Matthyssens, G., Kronenberger, P., Pays, E., Dero, B., Van Assel, S., Darville, M., Carvador, A., Steinert, M. & Hamers, R. (1983).** Gene activation and re-expression of a *Trypanosoma brucei* variant surface glycoprotein. *The EMBO Journal* **2** (7), 1185-1192.
- Miller, E. N. & Turner, M. J. (1981).** Analysis of antigenic types appearing in first relapse populations of clones of *Trypanosoma brucei*. *Parasitology* **82** 63-80.
- Miller, K. A., Sawicka, D., Barsky, D. & Albala, J. S. (2004).** Domain mapping of the Rad51 paralog protein complexes. *Nucleic Acids Research* **32** (1), 169-178.
- Mimitou, E. P. & Symington, L. S. (2008).** Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing. *Nature* **455** (7214), 770-774.
- Mims, C., Dimmock, N., Nash, A. & Stephen, J. (1995).** *Mim's pathogenesis of Infectious Disease*. London, UK: Academic Press.
- Miyazaki, J., Ryo, Y. & Minagawa, T. (1983).** Involvement of gene 49 in recombination of bacteriophage T4. *Genetics* **104** (1), 1-9.

- Miyazaki, T., Bressan, D. A., Shinohara, M., Haber, J. E. & Shinohara, A. (2004).** *in vivo* assembly and disassembly of Rad51 and Rad52 complexes during double-strand break repair. *The EMBO Journal* **23** 939-949.
- Morrison, H. G., McArthur, A. G., Gillin, F. D., Aley, S. B., Adam, R. D., Olsen, G. J., Best, A. A., Cande, W. Z., Chen, F., Cipriano, M. J., Davids, B. J., Dawson, S. C., Elmendorf, H. G., Hehl, A. B., Holder, M. E., Huse, S. M., Kim, U. U., Lasek-Nesselquist, E., Manning, G., Nigam, A., Nixon, J. E., Palm, D., Passamaneck, N. E., Prabhu, A., Reich, C. I., Reiner, D. S., Samuelson, J., Svard, S. G. & Sogin, M. L. (2007).** Genomic minimalism in the early diverging intestinal parasite *Giardia lamblia*. *Science* **317** (5846), 1921-1926.
- Morrison, L. J., Majiwa, P., Read, A. F. & Barry, J. D. (2005).** Probabilistic order in antigenic variation of *Trypanosoma brucei*. *International Journal for Parasitology* **35** (9), 961-972.
- Mortensen, U. H., Bendixen, C., Sunjevaric, I. & Rothstein, R. (1996).** DNA strand annealing is promoted by the yeast Rad52 protein. *Proceedings of the National Academy of Sciences* **93** (20), 10729-10734.
- Moynahan, M. E., Pierce, A. J. & Jasin, M. (2001).** BRCA2 is required for homology-directed repair of chromosomal breaks. *Cell* **7** (2), 263-272.
- Myler, P. J., Allison, J., Agabian, N. & Stuart, K. (1984).** Antigenic variation in African trypanosomes by gene replacement or activation of alternate telomeres. *Cell* **39** (1), 203-211.
- Nassif, N., Penney, J., Pal, S., Engels, W. R. & Gloor, G. B. (1994).** Efficient copying of nonhomologous sequences from ectopic sites via P-element-induced gap repair. *Molecular and Cellular Biology* **14** (3), 1613-1625.
- Navarro, M. & Gull, K. (2001).** A pol I transcriptional body associated with VSG mono-allelic expression in *Trypanosoma brucei*. *Nature* **414** (6865), 759-763.
- Neale, M. J., Pan, J. & Keeney, S. (2005).** Endonucleolytic processing of covalent protein-linked DNA double-strand breaks. *Nature* **436** (7053), 1053-1057.
- New, J. H., Sugiyama, T., Zaitseva, E. & Kowalczykowski, S. C. (1998).** Rad52 protein stimulates DNA strand exchange by Rad51 and replication protein A. *Nature* **391** (6665), 407-410.
- Norris, S. J. (2006).** Antigenic variation with a twist - the *Borrelia* story. *Molecular Microbiology* **60** (6), 1319-1322.
- O'Regan, P., Wilson, C., Townsend, S. & Thacker, J. (2001).** XRCC2 Is a nuclear RAD51-like protein required for damage-dependent RAD51 focus formation without the need for ATP-binding. *Journal of Biological Chemistry* **276** (25), 22148-22153.
- Oettinger, M. A., Schatz, D. G., Gorka, C. & Baltimore, D. (1990).** *RAG-1* and *RAG-2*, adjacent genes that synergistically activate V(D)J recombination. *Science* **248** (4962), 1517-1523.
- Ogawa, T., Yu, X., Shinohara, A. & Egelman, E. H. (1993).** Similarity of the yeast RAD51 filament to the bacterial RecA filament. *Science* **259** 1896-1899.

**Osakabe, K., Yoshioka, T., Ichikawa, H. & Toki, S. (2002).** Molecular cloning and characterization of RAD51-like genes from *Arabidopsis thaliana*. *Plant Molecular Biology* **50** (1), 71-81.

**Otero, R. D. C. & Hsieh, P. (1995).** Homologous recombination proteins in prokaryotes and eukaryotes. *Annual Review of Genetics* **29** (1), 509-552.

**Overath, P., Chaudhri, M., Steverding, D. & Ziegelbauer, K. (1994).** Invariant surface proteins in bloodstream forms of *Trypanosoma brucei*. *Parasitology Today* **10** (2), 53-58.

**Pain, A., Renauld, H., Berriman, M., Murphy, L., Yeats, C. A., Weir, W., Kerhornou, A., Aslett, M., Bishop, R., Bouchier, C., Cochet, M., Coulson, R. M., Cronin, A., de Villiers, E. P., Fraser, A., Fosker, N., Gardner, M., Goble, A., Griffiths-Jones, S., Harris, D. E., Katzer, F., Larke, N., Lord, A., Maser, P., McKellar, S., Mooney, P., Morton, F., Nene, V., O'Neil, S., Price, C., Quail, M. A., Rabinowitsch, E., Rawlings, N. D., Rutter, S., Saunders, D., Seeger, K., Shah, T., Squares, R., Squares, S., Tivey, A., Walker, A. R., Woodward, J., Dobbelaere, D. A., Langsley, G., Rajandream, M. A., McKeever, D., Shiels, B., Tait, A., Barrell, B. & Hall, N. (2005).** Genome of the host-cell transforming parasite *Theileria annulata* compared with *T. parva*. *Science* **309** (5731), 131-133.

**Palmer, G. H. & Brayton, K. A. (2007).** Gene conversion is a convergent strategy for pathogen antigenic variation. *Trends in Parasitology* **23** (9), 408-413.

**Paques, F. & Haber, J. E. (1999).** Multiple pathways of recombination induced by Double-Strand Breaks in *Saccharomyces cerevisiae*. *Microbiology and Molecular Biological Review* **63** (2), 349-404.

**Pays, E., Guyaux, M., Aerts, D., Van Meirvenne, N. & Steinert, M. (1985).** Telomeric reciprocal recombination as a possible mechanism for antigenic variation in trypanosomes. *Nature* **316** (6028), 562-564.

**Pays, E., Vanhamme, L. & Pérez-Morga, D. (2004).** Antigenic variation in *Trypanosoma brucei*: facts, challenges and mysteries. *Current Opinion in Microbiology* **7** (4), 369-374.

**Pays, E. (2006).** The variant surface glycoprotein as a tool for adaptation in African trypanosomes. *Microbes and Infection* **8** (3), 930-937.

**Pays, E., Vanhollebeke, B., Vanhamme, L., Paturiaux-Hanocq, F., Nolan, D. P. & Perez-Morga, D. (2006).** The trypanolytic factor of human serum. *Nature Reviews of Microbiology* **4** (6), 477-486.

**Pays, E., Salmon, D., Morrison, L. J., Marcello, L. & Barry, D. (2007).** Antigenic variation in *Trypanosoma brucei*. In *Trypanosomes: After the Genome*, pp. 339-372. Edited by D. Barry, R. McCulloch, J. C. Mottram & A. Acosta-Serrano. Norfolk, UK: Horizon Bioscience.

**Peacock, C. S., Seeger, K., Harris, D., Murphy, L., Ruiz, J. C., Quail, M. A., Peters, N., Adlem, E., Tivey, A., Aslett, M., Kerhornou, A., Ivens, A., Fraser, A., Rajandream, M. A., Carver, T., Norbertczak, H., Chillingworth, T., Hance, Z., Jagels, K., Moule, S., Ormond, D., Rutter, S., Squares, R., Whitehead, S., Rabinowitsch, E., Arrowsmith, C., White, B., Thurston, S., Bringaud, F., Baldauf, S. L., Faulconbridge,**

**A., Jeffares, D., Depledge, D. P., Oyola, S. O., Hilley, J. D., Brito, L. O., Tosi, L. R., Barrell, B., Cruz, A. K., Mottram, J. C., Smith, D. F. & Berriman, M. (2007).** Comparative genomic analysis of three *Leishmania* species that cause diverse human disease. *Nature Genetics* **39** (7), 839-847.

**Pellegrini, L., Yu, D. S., Lo, T., Anand, S., Lee, M., Blundell, T. L. & Venkitaraman, A. R. (2002).** Insights into DNA recombination from the structure of a RAD51-BRCA2 complex. *Nature* **420** (6913), 287-293.

**Petukhova, G., Stratton, S. & Sung, P. (1998).** Catalysis of homologous DNA pairing by yeast Rad51 and Rad54 proteins. *Nature* **393** (6680), 91-94.

**Pfeiffer, P., Thode, S., Hancke, J. & Vielmetter, W. (1994).** Mechanisms of overlap formation in nonhomologous DNA end joining. *Molecular and Cellular Biology* **14** (2), 888-895.

**Pierce, A. J., Johnson, R. D., Thompson, L. H. & Jasin, M. (1999).** XRCC3 promotes homology-directed repair of DNA damage in mammalian cells. *Genes & Development* **13** (20), 2633-2638.

**Pittman, D. L., Weinberg, L. R. & Schimenti, J. C. (1998).** Identification, Characterization, and Genetic Mapping of Rad51D, a New Mouse and Human RAD51/RecA-Related Gene. *Genomics* **49** (1), 103-111.

**Pittman, D. L. & Schimenti, J. C. (2000).** Midgestation lethality in mice deficient for the *RecA*-related gene, (*Rad51d/Rad51l3*). *Genesis* **26** (3), 167-173.

**Plug, A. W., Xu, J., Reddy, G., Golub, E. I. & Ashley, T. (1996).** Presynaptic association of Rad51 protein with selected sites in meiotic chromatin. *Proceedings of the National Academy of Sciences* **93** (12), 5920-5924.

**Pluskota-Karwatka, D. (2008).** Modifications of nucleosides by endogenous mutagens-DNA adducts arising from cellular processes. *Bioorganic Chemistry* **36** (4), 198-213.

**Pohl, T. J. & Nickoloff, J. A. (2008).** Rad51-independent interchromosomal Double-Strand Break repair by gene conversion requires Rad52 but not Rad55, Rad57, or Dmc1. *Molecular and Cellular Biology* **28** (3), 897-906.

**Proudfoot, C. (2005).** *RAD51*-like genes in *Trypanosoma brucei*: A potential role in antigenic variation., pp. 230: *PhD Thesis In: Wellcome Centre for Molecular Parasitology, University of Glasgow, Glasgow, UK.*

**Proudfoot, C. & McCulloch, R. (2005a).** Distinct roles for two *RAD51*-related genes in *Trypanosoma brucei* antigenic variation. *Nucleic Acids Research* **33** (21), 6906-6919.

**Proudfoot, C. & McCulloch, R. (2005b).** *Trypanosoma brucei* DMC1 does not act in DNA recombination, repair or antigenic variation in bloodstream stage cells. *Molecular and Biochemical Parasitology* **145** (2), 245-253.

**Puig, O., Caspary, F., Rigaut, G., Rutz, B., Bouveret, E., Bragado-Nilsson, E., Wilm, M. & Séraphin, B. (2001).** The Tandem Affinity Purification (TAP) method: A general procedure of protein complex purification. *Methods* **24** (3), 218-229.

- Raderschall, E., Golub, E. I. & Haaf, T. (1999).** Nuclear foci of mammalian recombination proteins are located at single-stranded DNA regions formed after DNA damage. *Progress in Nucleic Acid Research and Molecular Biology* **96** (5), 1921-1926.
- Ramsden, D. A. & Gellert, M. (1998).** Ku protein stimulates DNA end joining by mammalian DNA ligases: a direct role for Ku in repair of DNA double-strand breaks. *The EMBO Journal* **17** (2), 609-614.
- Register, J. C. & Griffith, J. (1985).** The direction of RecA protein assembly onto single strand DNA is the same as the direction of strand assimilation during strand exchange. *Journal of Biological Chemistry* **260** (22), 12308-12312.
- Richardson, C., Horikoshi, N. & Pandita, T. K. (2004).** The role of the DNA double-strand break response network in meiosis. *DNA Repair* **3** (8), 1149-1164.
- Rijkers, T., Van Den Ouweland, J., Morolli, B., Rolink, A. G., Baarends, W. M., Van Sloun, P. P. H., Lohman, P. H. M. & Pastink, A. (1998).** Targeted inactivation of mouse RAD52 reduces homologous recombination but not resistance to ionizing radiation. *Molecular and Cellular Biology* **18** (11), 6423-6429.
- Robertson, A., Klungland, A., Rognes, T. & Leiros, I. (2009).** Base excision repair: the long and short of it. *Cellular and Molecular Life Sciences*.
- Robinson, N. P., Burman, N., Melville, S. E. & Barry, J. D. (1999).** Predominance of duplicative VSG gene conversion in antigenic variation in African Trypanosomes. *Molecular and Cellular Biology* **19** (9), 5839-5846.
- Robinson, N. P., McCulloch, R., Conway, C., Browitt, A. & Barry, J. D. (2002).** Inactivation of MRE11 does not affect VSG gene duplication mediated by homologous recombination in *Trypanosoma brucei*. *Journal of Biological Chemistry* **277** (29), 26185-26193.
- Roditi, I. & Liniger, M. (2002).** Dressed for success: the surface coats of insect-borne protozoan parasites. *Trends in Microbiology* **10** (3), 128-134.
- Roeder, G. S. (1997).** Meiotic chromosomes: it takes two to tango. *Genes & Development* **11** (20), 2600-2621.
- Rolfes, R. J. (2006).** Regulation of purine nucleotide biosynthesis: in yeast and beyond. *Biochemical Society Transactions* **34** (5), 786-790.
- Roth, C., Longacre, S., Raibaud, A., Baltz, T. & Eisen, H. (1986).** The use of incomplete genes for the construction of a *Trypanosoma equiperdum* variant surface glycoprotein gene. *The EMBO Journal* **5** (5), 1065-1070.
- Roth, D. B. & Wilson, J. H. (1986).** Nonhomologous recombination in mammalian cells: role for short sequence homologies in the joining reaction. *Molecular and Cellular Biology* **6** (12), 4295-4304.
- Rubin, B. P., Ferguson, D. O. & Holloman, W. K. (1994).** Structure of REC2, a recombinational repair gene of *Ustilago maydis*, and its function in homologous recombination between plasmid and chromosomal sequences. *Molecular and Cellular Biology* **14** (9), 6287-6296.

- Rudenko, G., Blundell, P. A., Taylor, M. C., Kieft, R. & Borst, P. (1994).** VSG gene expression site control in insect form *Trypanosoma brucei*. *The EMBO Journal* **13** (22), 5470-5482.
- Rudenko, G., McCulloch, R., Dirks-Mulder, A. & Borst, P. (1996).** Telomere exchange can be an important mechanism of Variant Surface Glycoprotein gene switching in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology* **80** (1), 65-75.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989).** *Molecular Cloning - A Laboratory Manual*, Second edn. New York, USA.: Cold Spring Harbour Laboratory Press.
- San Filippo, J., Sung, P. & Klein, H. (2008).** Mechanism of eukaryotic homologous recombination. *Annual Review of Biochemistry* **77** (1).
- Sandler, S. J., Satin, L. H., Samra, H. S. & Clark, A. J. (1996).** recA-like genes from three archaean species with putative protein products similar to Rad51 and Dmc1 proteins of the yeast *Saccharomyces cerevisiae*. *Nucleic Acids Research* **24** (11), 2125-2132.
- Sather, S. & Agabian, N. (1985).** A 5' spliced leader is added in trans to both alpha- and beta-tubulin transcripts in *Trypanosoma brucei*. *Proceedings of the National Academy of Sciences* **82** (17), 5695-5699.
- Schatz, D. G., Oettinger, M. A. & Baltimore, D. (1989).** The V(D)J recombination activating gene, *RAG-1*. *Cell* **59** (6), 1035-1048.
- Schell, D., Evers, R., Preis, D., Ziegelbauer, K., Kiefer, H., Lottspeich, F., Cornelissen, A. W. & Overath, P. (1991).** A transferrin-binding protein of *Trypanosoma brucei* is encoded by one of the genes in the variant surface glycoprotein gene expression site. *The EMBO Journal* **10** 1061-1066.
- Schild, D., Lio, Y., Collins, D. W., Tsomondo, T. & Chen, D. J. (2000).** Evidence for simultaneous protein interactions between human Rad51 paralogs. *Journal of Biological Chemistry* **275** (22), 16443-16449.
- Schimanski, B., Nguyen, T. N. & Gunzl, A. (2005a).** Highly efficient Tandem Affinity Purification of trypanosome protein complexes based on a novel epitope combination. *Eukaryotic Cell* **4** (11), 1942-1950.
- Schimanski, B., Nguyen, T. N. & Gunzl, A. (2005b).** Characterization of a multi-subunit transcription factor complex essential for Spliced-Leader RNA gene transcription in *Trypanosoma brucei*. *Molecular and Cellular Biology* **25** (16), 7303-7313.
- Schmidt, G. D. & Roberts, L. S. (1989).** Foundations of Parasitology. pp. 55-80. Edited by D. K. Brake. New York, USA: McGraw Hill International Edition.
- Schuermann, D., Molinier, J., Fritsch, O. & Hohn, B. (2005).** The dual nature of homologous recombination in plants. *Trends in Genetics* **21** (3), 172-181.
- Scully, R., Chen, J., Plug, A., Xiao, Y., Weaver, D., Feunteun, J., Ashley, T. & Livingston, D. M. (1997).** Association of Brca1 with Rad51 in mitotic and meiotic cells. *Cell* **88** (2), 265-275.
- Seed, J. & Wenck, M. (2003).** Role of the long slender to short stumpy transition in the life cycle of the african trypanosomes. *Kinetoplastid Biology and Disease* **2** (1), 3.



**SGD-project (*Saccharomyces Genome Database*).**2008: <ftp://ftp.yeastgenome.org/yeast/>.

**Shah, J. S., Young, J. R., Kimmel, B. E., Iams, K. P. & Williams, R. O. (1987).** The 5' flanking sequence of a *Trypanosoma brucei* variable surface glycoprotein gene. *Molecular and Biochemical Parasitology* **24** (2), 163-174.

**Shao, X. & Grishin, N. V. (2000).** Common fold in helix-hairpin-helix proteins. *Nucleic Acids Research* **28** (14), 2643-2650.

**Shea, C., Glass, D. J., Parangi, S. & Van der Ploeg, L. H. (1986).** Variant surface glycoprotein gene expression site switches in *Trypanosoma brucei*. *Journal of Biological Chemistry* **261** (13), 6056-6063.

**Shim, K. S., Schmutte, C., Tomblin, G., Heinen, C. D. & Fishel, R. (2004).** hXRCC2 enhances ADP/ATP processing and strand exchange by hRAD51. *Journal of Biological Chemistry* **279** (29), 30385-30394.

**Shin, D. S., Pellegrini, L., Daniels, D. S., Yelent, B., Craig, L., Bates, D., Yu, D. S., Shivji, M. K., Hitomi, C., Arvai, A. S., Volkmann, N., Tsuruta, H., Blundell, T. L., Venkitaraman, A. R. & Tainer, J. A. (2003).** Full-length archaeal Rad51 structure and mutants: mechanisms for RAD51 assembly and control by BRCA2. *The EMBO Journal* **22** (17), 4566-4576.

**Shinohara, A., Ogawa, H., Matsuda, Y., Ushio, N., Ikeo, K. & Ogawa, T. (1993).** Cloning of human, mouse and fission yeast recombination genes homologous to *RAD51* and *RecA*. *Nature Genetics* **4** (3), 239-243.

**Shinohara, A. & Ogawa, T. (1998).** Stimulation by Rad52 of yeast Rad51-mediated recombination. *Nature* **391** (6665), 404-407.

**Shinohara, M., Gasior, S. L., Bishop, D. K. & Shinohara, A. (2000).** Tid1/Rdh54 promotes colocalization of Rad51 and Dmc1 during meiotic recombination. *Proceedings of the National Academy of Sciences* **97** (20), 10814-10819.

**Shu, Z., Smith, S., Wang, L., Rice, M. C. & Kmiec, E. B. (1999).** Disruption of *muREC2/RAD51L1* in mice results in early embryonic lethality which can be partially rescued in a *p53*<sup>-/-</sup> background. *Molecular and Cellular Biology* **19** (12), 8686-8693.

**Sibley, L. D. (2004).** Intracellular parasite invasion strategies. *Science* **304** (5668), 248-253.

**Signon, L., Malkova, A., Naylor, M. L., Klein, H. & Haber, J. E. (2001).** Genetic requirements for RAD51- and RAD54-independent break-induced replication repair of a chromosomal Double-Strand Break. *Molecular and Cellular Biology* **21** (6), 2048-2056.

**Sigurdsson, S., Van Komen, S., Bussen, W., Schild, D., Albala, J. S. & Sung, P. (2001).** Mediator function of the human Rad51B-Rad51C complex in Rad51/RPA-catalyzed DNA strand exchange. *Genes and Development* **15** (24), 3308-3318.

**Smeaton, M. B., Miller, P. S., Ketner, G. & Hanakahi, L. A. (2007).** Small-scale extracts for the study of nucleotide excision repair and non-homologous end joining. *Nucleic Acids Research* **35** (22), e152.

- Smiraldo, P. G., Gruver, A. M., Osborn, J. C. & Pittman, D. L. (2005).** Extensive chromosomal instability in *Rad51d*-deficient mouse cells. *Cancer Research* **65** (6), 2089-2096.
- Sommer, J. M., Cheng, Q. L., Keller, G. A. & Wang, C. C. (1992).** In vivo import of firefly luciferase into the glycosomes of *Trypanosoma brucei* and mutational analysis of the C-terminal targeting signal. *Molecular Biology of the Cell* **3** (7), 749-759.
- Sonoda, E., Sasaki, M. S., Buerstedde, J. M., Bezzubova, O., Shinohara, A., Ogawa, H., Takata, M., Yamaguchi-Iwai, Y. & Takeda, S. (1998).** Rad51-deficient vertebrate cells accumulate chromosomal breaks prior to cell death. *The EMBO Journal* **17** (2), 598-608.
- Southern, E. M. (1975).** Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology* **98** (3), 503-517.
- Stassen, N. Y., Logsdon Jr, J. M., Vora, G. J., Offenberg, H. H., Palmer, J. D. & Zolan, M. E. (1997).** Isolation and characterization of rad51 orthologs from *Coprinus cinereus* and *Lycopersicon esculentum*, and phylogenetic analysis of eukaryotic recA homologs. *Current Genetics* **31** (2), 144-157.
- Sternberg, J. M. (2004).** Human African trypanosomiasis: clinical presentation and immune response. *Parasite Immunology* **26** (11), 469-476.
- Steverding, D., Stierhof, Y. D., Fuchs, H., Tauber, R. & Overath, P. (1995).** Transferrin-binding protein complex is the receptor for transferrin uptake in *Trypanosoma brucei*. *Journal of Cellular Biology* **131** (5), 1173-1182.
- Steverding, D. (2008).** The history of African trypanosomiasis. *Parasites & Vectors* **1** (1), 3.
- Stich, A., Abel, P. M. & Krishna, S. (2002).** Human African trypanosomiasis. *British Medical Journal* **325** (7357), 203-206.
- Stockdale, C., Swiderski, M. R., Barry, J. D. & McCulloch, R. (2008).** Antigenic variation in *Trypanosoma brucei*: joining the DOTs. *Public Library of Science, Biology* **6** (7), 13861391.
- Story, R. M., Weber, I. T. & Steitz, T. A. (1992).** The structure of the *E. coli* recA protein monomer and polymer. *Nature* **355** (6358), 318-325.
- Story, R. M., Bishop, D. K., Kleckner, N. & Steitz, T. A. (1993).** Structural relationship of bacterial RecA proteins to recombination proteins from bacteriophage T4 and yeast. *Science* **259** (5103), 1892-1896.
- Strathern, J. N., Klar, A. J. S., Hicks, J. B., Abraham, J. A., Ivy, J. M., Nasmyth, K. A. & McGill, C. (1982).** Homothallic switching of yeast mating type cassettes is initiated by a double-stranded cut in the MAT locus. *Cell* **31** (1), 183-192.
- Sugawara, N. & Haber, J. E. (1992).** Characterization of double-strand break-induced recombination: homology requirements and single-stranded DNA formation. *Molecular and Cellular Biology* **12** (2), 563-575.

- Sugawara, N., Ivanov, E. L., Fishman-Lobell, J., Ray, B. L., Wu, X. & Haber, J. E. (1995).** DNA structure-dependent requirements for yeast *RAD* genes in gene conversion. *Nature* **373** (5), 84-86.
- Sugawara, N., Wang, X. & Haber, J. E. (2003).** *in vivo* roles of Rad52, Rad54, and Rad55 proteins in Rad51-mediated recombination. *Cell* **12** (1), 209-219.
- Sugiyama, T., New, J. H. & Kowalczykowski, S. C. (1998).** DNA annealing by Rad52 Protein is stimulated by specific interaction with the complex of replication protein A and single-stranded DNA. *Proceedings of the National Academy of Sciences* **95** (11), 6049-6054.
- Sung, P. & Robberson, D. L. (1995).** DNA strand exchange mediated by a RAD51-ssDNA nucleoprotein filament with polarity opposite to that of RecA. *Cell* **82** (3), 453-461.
- Sung, P. (1997).** Yeast Rad55 and Rad57 proteins form a heterodimer that functions with replication protein A to promote DNA strand exchange by Rad51 recombinase. *Genes & Development* **11** (9), 1111-1121.
- Sung, P., Trujillo, K. M. & Van Komen, S. (2000).** Recombination factors of *Saccharomyces cerevisiae*. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **451** (1-2), 257-275.
- Swoboda, P., Gal, S., Hohn, B. & Puchta, H. (1994).** Intrachromosomal homologous recombination in whole plants. *The EMBO Journal* **13** (2), 484-489.
- Symington, L. S. (2002).** Role of RAD52 epistasis group genes in homologous recombination and double-strand break repair. *Microbiology and Molecular Biology Reviews* **66** (4), 630-670.
- Szostak, J. W., Orr-Weaver, T. L., Rothstein, R. J. & Stahl, F. W. (1983).** The double-strand-break repair model for recombination. *Cell* **33** (1), 25-35.
- Takanami, T., Sato, S., Ishihara, T., Katsura, I., Takahashi, H. & Higashitani, A. (1998).** Characterization of a *Caenorhabditis elegans* *recA*-like gene *Ce-rdh-1* involved in meiotic recombination. *DNA Research* **5** (6), 373-377.
- Takata, M., Sasaki, M. S., Sonoda, E., Fukushima, T., Morrison, C., Albala, J. S., Swagemakers, S. M. A., Kanaar, R., Thompson, L. H. & Takeda, S. (2000).** The Rad51 paralog Rad51B promotes homologous recombinational repair. *Molecular and Cellular Biology* **20** (17), 6476-6482.
- Takata, M., Sasaki, M. S., Tachiiri, S., Fukushima, T., Sonoda, E., Schild, D., Thompson, L. H. & Takeda, S. (2001).** Chromosome instability and defective recombinational repair in knockout mutants of the five Rad51 paralogs. *Molecular and Cellular Biology* **21** (8), 2858-2866.
- Tarsounas, M., Davies, A. A. & West, S. C. (2004a).** RAD51 localization and activation following DNA damage. *Philosophical Transaction of the Royal Society, London* **359** 87-93.
- Tarsounas, M., Muñoz, P. O., Claas, A., Smiraldo, P. G., Pittman, D. L., Blasco, M. A. & West, S. C. (2004b).** Telomere maintenance requires the RAD51D recombination/repair protein. *Cell* **117** (3), 337-347.

- Tashiro, S., Walter, J., Shinohara, A., Kamada, N. & Cremer, T. (2000).** Rad51 accumulation at sites of DNA damage and in postreplicative chromatin. *Journal of Cell Biology* **150** (2), 283-292.
- Tebbs, R. S., Zhao, Y., Tucker, J. D., Scheerer, J. B., Siciliano, M. J., Hwang, M., Liu, N., Legerski, R. J. & Thompson, L. H. (1995).** Correction of chromosomal instability and sensitivity to diverse mutagens by a cloned cDNA of the XRCC3 DNA repair gene. *Proceedings of the National Academy of Sciences* **92** (14), 6354-6358.
- Thompson, L. H. & Schild, D. (2001).** Homologous recombinational repair of DNA ensures mammalian chromosome stability. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **477** (1), 131-153.
- Tonegawa, S. (1983).** Somatic generation of antibody diversity. *Nature* **302** (5909), 575-581.
- Towbin, H. & Gordon, J. (1984).** Immunoblotting and dot immunobinding - Current status and outlook. *Journal of Immunological Methods* **72** (2), 313-340.
- Tsutsui, Y., Khasanov, F. K., Shinagawa, H., Iwasaki, H. & Bashkirov, V. I. (2001).** Multiple interactions among the components of the recombinational DNA repair system in *Schizosaccharomyces pombe*. *Genetics* **159** (1), 91-105.
- Tsuzuki, T., Fujii, Y., Sakumi, K., Tominaga, Y., Nakao, K., Sekiguchi, M., Matsushiro, A., Yoshimura, Y. & Morita T (1996).** Targeted disruption of the *Rad51* gene leads to lethality in embryonic mice. *Proceedings of the National Academy of Sciences* **93** (13), 6236-6240.
- Turner, C. M., Barry, J. D., Maudlin, I. & Vickerman, K. (1988).** An estimate of the size of the metacyclic variable antigen repertoire of *Trypanosoma brucei rhodesiense*. *Parasitology* **97** (2), 269-276.
- Turner, C. M., Sternberg, J., Buchanan, N., Smith, E., Hide, G. & Tait, A. (1990).** Evidence that the mechanism of gene exchange in *Trypanosoma brucei* involves meiosis and syngamy. *Parasitology* **101** (3), 377-386.
- Turner, C. M., McLellan, S., Lindergard, L. A., Bisoni, L., Tait, A. & MacLeod, A. (2004).** Human infectivity trait in *Trypanosoma brucei*: stability, heritability and relationship to sra expression. *Parasitology* **129** (4), 445-454.
- Ugolini, S. & Bruschi, C. V. (1996).** The red/white colony color assay in the yeast *Saccharomyces cerevisiae*: epistatic growth advantage of white *ade8-18*, *ade2* cells over red *ade2* cells. *Current Genetics* **30** (6), 485-492.
- Ullu, E., Matthews, K. R. & Tschudi, C. (1993).** Temporal order of RNA-processing reactions in trypanosomes: rapid *trans* splicing precedes polyadenylation of newly synthesized tubulin transcripts. *Molecular and Cellular Biology* **13** (1), 720-725.
- van der Heijden, T., Seidel, R., Modesti, M., Kanaar, R., Wyman, C. & Dekker, C. (2007).** Real-time assembly and disassembly of human RAD51 filaments on individual DNA molecules. *Nucleic Acids Research* **35** (17), 5646-5657.

**van der Woude, M. W. & Baumler, A. J. (2004).** Phase and Antigenic Variation in Bacteria. *Clinical Microbiology Reviews* **17** (3), 581-611.

**van Veelen, L. R., Essers, J., van de Rakt, M. W. M., Odijk, H., Pastink, A., Zdzienicka, M. Z., Paulusma, C. C. & Kanaar, R. (2005).** Ionizing radiation-induced foci formation of mammalian Rad51 and Rad54 depends on the Rad51 paralogs, but not on Rad52. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **574** 34-49.

**Van Xong, H., Vanhamme, L., Chamekh, M., Chimfwembe, C. E., Van Den Abbeele, J., Pays, A., Van Meirvenne, N., Hamers, R., De Baetselier, P. & Pays, E. (1998).** A VSG Expression Site-Associated Gene confers resistance to human serum in *Trypanosoma rhodesiense*. *Cell* **95** (6), 839-846.

**Vickerman, K. (1969).** On the surface coat and flagellar adhesion in trypanosomes. *Journal of Cell Science* **5** (1), 163-193.

**Vickerman, K. (1985).** Developmental cycles and biology of pathogenic trypanosomes. *British Medical Bulletin* **41** (2), 105-114.

**Viviani, V. R. (2002).** The origin, diversity, and structure function relationships of insect luciferases. *Cellular and Molecular Life Sciences* **59** (11), 1833-1850.

**Vojtek, A. B., Hollenberg, S. M. & Cooper, J. A. (1993).** Mammalian Ras interacts directly with the serine/threonine kinase raf. *Cell* **74** (1), 205-214.

**Walker, J. E., Saraste, M., Runswick, M. J. & Gay, N. J. (1982).** Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *The EMBO Journal* **1** (8), 945-951.

**Wang, H., Perrault, A. P., Takeda, Y., Qin, W., Wang, H. & Lliakis, G. (2003).** Biochemical evidence for Ku-independent back-up pathways of NHEJ. *Nucleic Acids Research* **31** (18), 5377-5388.

**Wang, X., Ira, G., Tercero, J. A., Holmes, A. M., Diffley, J. F. X. & Haber, J. E. (2004).** Role of DNA replication proteins in Double-Strand Break-induced recombination in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology* **24** (16), 6891-6899.

**Ward, J. D., Barber, L. J., Petalcorin, M. I., Yanowitz, J. & Boulton, S. J. (2007).** Replication blocking lesions present a unique substrate for homologous recombination. *The EMBO Journal* **24** (14), 3384-3396.

**Weiffenbach, B. & Haber, J. E. (1981).** Homothallic mating type switching generates lethal chromosome breaks in *rad52* strains of *Saccharomyces cerevisiae*. *Molecular and Cellular Biology* **1** (6), 522-534.

**Weterings, E. & van Gent, D. C. (2004).** The mechanism of non-homologous end-joining a synopsis of synapsis. *DNA Repair* **3** 1425-1435.

**Weterings, E. & Chen, D. J. (2007).** DNA-dependent protein kinase in nonhomologous end joining: a lock with multiple keys? *Journal of Cell Biology* **179** (2), 183-186.

**White, C. I. & Haber, J. E. (1990).** Intermediates of recombination during mating type switching in *Saccharomyces cerevisiae*. *The EMBO Journal* **9** (3), 663-673.

**WHO (2009).** <http://www.who.int/en/>.

**Wickstead, B., Ersfeld, K. & Gull, K. (2004).** The small chromosomes of *Trypanosoma brucei* involved in antigenic variation are constructed around repetitive palindromes. *Genome Research* **14** (6), 1014-1024.

**Wiese, C., Collins, D., Albala, J., Thompson, L., Kronenberg, A. & Schild, D. (2002).** Interactions involving the Rad51 paralogs Rad51C and XRCC3 in human cells. *Nucleic Acids Research* **30** (4), 1001-1008.

**Wiese, C., Hinz, J. M., Tebbs, R. S., Nham, P. B., Urbin, S. S., Collins, D. W., Thompson, L. H. & Schild, D. (2006).** Disparate requirements for the Walker A and B ATPase motifs of human RAD51D in homologous recombination. *Nucleic Acids Research* **34** (9), 2833-2843.

**Wilson, J. W., Schurr, M. J., LeBlanc, C. L., Ramamurthy, R., Buchanan, K. L. & Nickerson, C. A. (2002).** Mechanisms of bacterial pathogenicity. *PMJ online* **78** (918), 216-224.

**Woods, W. G. & Dyall-Smith, M. L. (1997).** Construction and analysis of a recombination-deficient (*radA*) mutant of *Haloferax volcanii*. *Molecular Microbiology* **23** (4), 791-797.

**Wyatt, M. D. & Pittman, D. L. (2006).** Methylating agents and DNA repair responses: methylated bases and sources of strand breaks. *Chemical research in toxicology* **19** (12), 1580-1594.

**Wyman, C., Ristic, D. & Kanaar, R. (2004).** Homologous recombination-mediated double-strand break repair. *DNA Repair* **3** (8), 827-833.

**Yamada, N. A., Hinz, J. M., Kopf, V. L., Segalle, K. D. & Thompson, L. H. (2004).** XRCC3 ATPase activity is required for normal XRCC3-Rad51C complex dynamics and homologous recombination. *Journal of Biological Chemistry* **279** (22), 23250-23254.

**Yamaguchi-Iwai, Y., Sonoda, E., Buerstedde, J.-M., Bezzubova, O., Morrison, C., Takata, M., Shinohara, A. & Takeda, S. (1998).** Homologous recombination, but not DNA repair, is reduced in vertebrate cells deficient in RAD52. *Molecular and Cellular Biology* **18** (11), 6430-6435.

**Yan, C. T., Boboila, C., Souza, E. K., Franco, S., Hickernell, T. R., Murphy, M., Gumaste, S., Geyer, M., Zarrin, A. A., Manis, J. P., Rajewsky, K. & Alt, F. W. (2007).** IgH class switching and translocations use a robust non-classical end-joining pathway. *Nature* **449** (7161), 478-482.

**Yaneva, M., Kowalewski, T. & Lieber, M. R. (1997).** Interaction of DNA-dependent protein kinase with DNA and with Ku: biochemical and atomic-force microscopy studies. *The EMBO Journal* **16** (16), 5098-5112.

**Yang, H., Jeffrey, P. D., Miller, J., Kinnucan, E., Sun, Y., Thoma, N. H., Zheng, N., Chen, P. L., Lee, W. H. & Pavletich, N. P. (2002).** BRCA2 function in DNA binding and recombination from a BRCA2-DSS1-ssDNA structure. *Science* **297** (5588), 1837-1848.

- Yanowitz, J. L. (2008).** Genome integrity is regulated by the *Caenorhabditis elegans* *Rad51D* homolog *rfs-1*. *Genetics* **179** (1), 249-262.
- Yonesaki, T., Ryo, Y., Minagawa, T. & Takahashi, H. (1985).** Purification and some of the functions of the products of bacteriophage T4 recombination genes, *uvs X* and *uvs Y*. *European Journal of Biochemistry* **148** (1), 127-134.
- Young, R., Taylor, J., Kurioka, A., Becker, M., Louis, E. & Rudenko, G. (2008).** Isolation and analysis of the genetic diversity of repertoires of VSG expression site containing telomeres from *Trypanosoma brucei gambiense*, *T. b. brucei* and *T. equiperdum*. *BMC Genomics* **9** (1), 385.
- Zambrano-Villa, S., Rosales-Borjas, D., Carrero, J. C. & Ortiz-Ortiz, L. (2002).** How protozoan parasites evade the immune response. *Trends in Parasitology* **18** (6), 272-278.
- Zhang, J.-R., Hardham, J. M., Barbour, A. G. & Norris, S. J. (1997).** Antigenic variation in Lyme Disease *Borreliae* by promiscuous recombination of VMP-like sequence cassettes. *Cell* **89** (2), 275-285.
- Zhang, Z., Fan, H.-Y., Goldman, J. A. & Kingston, R. E. (2007).** Homology-driven chromatin remodeling by human RAD54. *Nature Structural & Molecular Biology* **14** (5), 397-405.